



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification 6 : C12N 15/12, 5/10, 1/21, C07K 14/47, 16/18, C12N 1/21, C07K 14/47, 16/18, C12Q 1/68, G01N 33/50, 33/53, 33/68, A61K 38/17</p>	<p>A2</p>	<p>(11) International Publication Number: WO 98/39448 (43) International Publication Date: 11 September 1998 (11.09.98)</p>																														
<p>(21) International Application Number: PCT/US98/04493 (22) International Filing Date: 6 March 1998 (06.03.98)</p> <p>(30) Priority Data:</p> <table border="0"> <tr><td>60/040,162</td><td>7 March 1997 (07.03.97)</td><td>US</td></tr> <tr><td>60/040,333</td><td>7 March 1997 (07.03.97)</td><td>US</td></tr> <tr><td>60/038,621</td><td>7 March 1997 (07.03.97)</td><td>US</td></tr> <tr><td>60/040,161</td><td>7 March 1997 (07.03.97)</td><td>US</td></tr> <tr><td>60/040,626</td><td>7 March 1997 (07.03.97)</td><td>US</td></tr> <tr><td>60/040,334</td><td>7 March 1997 (07.03.97)</td><td>US</td></tr> <tr><td>60/040,336</td><td>7 March 1997 (07.03.97)</td><td>US</td></tr> <tr><td>60/040,163</td><td>7 March 1997 (07.03.97)</td><td>US</td></tr> <tr><td>60/043,580</td><td>11 April 1997 (11.04.97)</td><td>US</td></tr> <tr><td>60/043,568</td><td>11 April 1997 (11.04.97)</td><td>US</td></tr> </table> <p>(Continued on the following page)</p> <p>(71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).</p> <p>(72) Inventors; and (75) Inventors/Applicants (for US only): RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hills Road, Laytonville, MD 20882 (US). FISCHER, Carrie, L. [US/US]; 5810 Hall Street, Burke, VA 22015 (US). SOPENET, Daniel, R. [US/US]; 15050 Stillfield, Place, Centreville, VA 22020 (US). CARTER, Kenneth, C. [US/US]; 11601 Brandy Hall Lane, North Potomac, MD 20878 (US). BEDNARIK, Daniel, P. [US/US]; 8822 Blue Sea Drive, Columbia, MD 21046 (US). ENDRESS, Gregory, A. [US/US]; 9729 Clagett Farm Drive, Potomac, MD 20854 (US). YU, Guo-Liang [CN/US]; 13524 Straw Bale Lane, Darnestown, MD 20878 (US). NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). FENG, Ping [CN/US]; 4 Relda Court, Gaithersburg, MD 20878 (US). YOUNG, Paul, E. [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). GREENE, John, M. [US/US]; 872 Diamond Drive, Gaithersburg, MD 20878 (US). FERRIE, Ann, M. [US/US]; 13203 L Astoria Hill Court, Germantown,</p>		60/040,162	7 March 1997 (07.03.97)	US	60/040,333	7 March 1997 (07.03.97)	US	60/038,621	7 March 1997 (07.03.97)	US	60/040,161	7 March 1997 (07.03.97)	US	60/040,626	7 March 1997 (07.03.97)	US	60/040,334	7 March 1997 (07.03.97)	US	60/040,336	7 March 1997 (07.03.97)	US	60/040,163	7 March 1997 (07.03.97)	US	60/043,580	11 April 1997 (11.04.97)	US	60/043,568	11 April 1997 (11.04.97)	US	<p>MD 20874 (US). DUAN, Roxanne [US/US]; 4541 Fairfield Drive, Bethesda, MD 20814 (US). HU, Jing-Shan [CN/US]; 1247 Lakeside Drive #3034, Sunnyvale, CA 94086 (US). FLORENCE, Kimberly, A. [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick Place #24, Gaithersburg, MD 20878 (US). EBNER, Reinhard [DE/US]; 9906 Shelburne Terrace #316, Gaithersburg, MD 20878 (US). BREWER, Laurie, A. [US/US]; 14920 Mount Nebo Road, Poolesville, MD 20837 (US). MOORE, Paul, A. [GB/US]; Apartment #104, 1908 Holly Ridge Drive, McLean, VA 22102 (US). SHI, Yanggu [CN/US]; 437 West Side Drive, Gaithersburg, MD 20878 (US). LAFLEUR, David, W. [US/US]; 1615 Q Street, N.W. #807, Washington, DC 20009 (US). LI, Yi [CN/US]; 1247 Lakeside Drive #3034, Sunnyvale, CA 94086 (US). ZENG, Zhizhen [CN/US]; 13950 Saddleview Drive, Gaithersburg, MD 20878 (US). KYAW, Hla [BU/US]; 520 Sugarbush Circle, Frederick, MD 21703 (US).</p> <p>(74) Agents: BROOKES, Anders, A. et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 10850 (US).</p> <p>(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</p> <p>Published Without international search report and to be republished upon receipt of that report.</p>
60/040,162	7 March 1997 (07.03.97)	US																														
60/040,333	7 March 1997 (07.03.97)	US																														
60/038,621	7 March 1997 (07.03.97)	US																														
60/040,161	7 March 1997 (07.03.97)	US																														
60/040,626	7 March 1997 (07.03.97)	US																														
60/040,334	7 March 1997 (07.03.97)	US																														
60/040,336	7 March 1997 (07.03.97)	US																														
60/040,163	7 March 1997 (07.03.97)	US																														
60/043,580	11 April 1997 (11.04.97)	US																														
60/043,568	11 April 1997 (11.04.97)	US																														
<p>(54) Title: 186 HUMAN SECRETED PROTEINS</p> <p>(57) Abstract</p> <p>The present invention relates to 186 novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.</p>																																

BEST AVAILABLE COPY

60/043,314	11 april 1997 (11.04.97)	US	60/047,596	23 May 1997 (23.05.97)	US	60/056,637	22 August 1997 (22.08.97)	US
60/043,569	11 april 1997 (11.04.97)	US	60/047,612	23 May 1997 (23.05.97)	US	60/056,903	22 August 1997 (22.08.97)	US
60/043,311	11 april 1997 (11.04.97)	US	60/047,632	23 May 1997 (23.05.97)	US	60/056,888	22 August 1997 (22.08.97)	US
60/043,671	11 april 1997 (11.04.97)	US	60/047,601	23 May 1997 (23.05.97)	US	60/056,879	22 August 1997 (22.08.97)	US
60/043,674	11 april 1997 (11.04.97)	US	60/047,595	23 May 1997 (23.05.97)	US	60/056,880	22 August 1997 (22.08.97)	US
60/043,669	11 april 1997 (11.04.97)	US	60/047,599	23 May 1997 (23.05.97)	US	60/056,894	22 August 1997 (22.08.97)	US
60/043,312	11 april 1997 (11.04.97)	US	60/047,588	23 May 1997 (23.05.97)	US	60/056,911	22 August 1997 (22.08.97)	US
60/043,313	11 april 1997 (11.04.97)	US	60/047,585	23 May 1997 (23.05.97)	US	60/056,636	22 August 1997 (22.08.97)	US
60/043,672	11 april 1997 (11.04.97)	US	60/047,586	23 May 1997 (23.05.97)	US	60/056,874	22 August 1997 (22.08.97)	US
60/043,578	11 april 1997 (11.04.97)	US	60/047,590	23 May 1997 (23.05.97)	US	60/056,910	22 August 1997 (22.08.97)	US
60/043,576	11 april 1997 (11.04.97)	US	60/047,594	23 May 1997 (23.05.97)	US	60/056,864	22 August 1997 (22.08.97)	US
60/043,670	11 april 1997 (11.04.97)	US	60/047,589	23 May 1997 (23.05.97)	US	60/056,631	22 August 1997 (22.08.97)	US
60/047,600	23 May 1997 (23.05.97)	US	60/047,593	23 May 1997 (23.05.97)	US	60/056,845	22 August 1997 (22.08.97)	US
60/047,615	23 May 1997 (23.05.97)	US	60/047,614	23 May 1997 (23.05.97)	US	60/056,892	22 August 1997 (22.08.97)	US
60/047,597	23 May 1997 (23.05.97)	US	60/047,501	23 May 1997 (23.05.97)	US	60/056,632	22 August 1997 (22.08.97)	US
60/047,502	23 May 1997 (23.05.97)	US	60/048,974	06 June 1997 (06.06.97)	US	60/056,664	22 August 1997 (22.08.97)	US
60/047,633	23 May 1997 (23.05.97)	US	60/048,964	06 June 1997 (06.06.97)	US	60/056,876	22 August 1997 (22.08.97)	US
60/047,583	23 May 1997 (23.05.97)	US	60/049,610	13 June 1997 (13.06.97)	US	60/056,881	22 August 1997 (22.08.97)	US
60/047,617	23 May 1997 (23.05.97)	US	60/051,926	08 July 1997 (08.07.97)	US	60/056,909	22 August 1997 (22.08.97)	US
60/047,618	23 May 1997 (23.05.97)	US	60/052,874	16 July 1997 (16.07.97)	US	60/056,875	22 August 1997 (22.08.97)	US
60/047,503	23 May 1997 (23.05.97)	US	60/055,724	18 August 1997 (18.08.97)	US	60/056,862	22 August 1997 (22.08.97)	US
60/047,592	23 May 1997 (23.05.97)	US	60/056,886	22 August 1997 (22.08.97)	US	60/056,887	22 August 1997 (22.08.97)	US
60/047,581	23 May 1997 (23.05.97)	US	60/056,877	22 August 1997 (22.08.97)	US	60/056,908	22 August 1997 (22.08.97)	US
60/047,584	23 May 1997 (23.05.97)	US	60/056,889	22 August 1997 (22.08.97)	US	60/056,884	22 August 1997 (22.08.97)	US
60/047,500	23 May 1997 (23.05.97)	US	60/056,893	22 August 1997 (22.08.97)	US	60/057,761	05 September 1997 (05.09.97)	US
60/047,587	23 May 1997 (23.05.97)	US	60/056,630	22 August 1997 (22.08.97)	US	60/057,650	05 September 1997 (05.09.97)	US
60/047,492	23 May 1997 (23.05.97)	US	60/056,878	22 August 1997 (22.08.97)	US	60/057,669	05 September 1997 (05.09.97)	US
60/047,598	23 May 1997 (23.05.97)	US	60/056,662	22 August 1997 (22.08.97)	US	60/058,785	12 September 1997 (12.09.97)	US
60/047,613	23 May 1997 (23.05.97)	US	60/056,872	22 August 1997 (22.08.97)	US	60/061,060	02 October 1997 (02.10.97)	US
60/047,582	23 May 1997 (23.05.97)	US	60/056,882	22 August 1997 (22.08.97)	US			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

186 Human Secreted Proteins

Field of the Invention

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and
5 their production.

Background of the Invention

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or
10 organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum
15 (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

20 Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or
25 secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include
30 the commercially valuable human insulin, interferon, Factor VIII, human growth hormone, tissue plasminogen activator, and erythropoietin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using
35 secreted proteins or the genes that encode them.

Summary of the Invention

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

Detailed Description

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig

analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 12301 Park Lawn Drive, Rockville, Maryland 20852, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA contained within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH₂PO₄; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA⁺ sequences (such as any 3' terminal polyA⁺ tract of a cDNA shown in the sequence listing), or to a

complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

5 The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be
10 single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability
15 or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

 The polypeptide of the present invention can be composed of amino acids joined
20 to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs,
25 as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be
30 branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a
35 nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine,

formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins
5 such as arginylation, and ubiquitination. (See, for instance, PROTEINS -
STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W.
H. Freeman and Company, New York (1993); POSTTRANSLATIONAL
COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic
Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990);
10 Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting
15 activity similar, but not necessarily identical to, an activity of a polypeptide of the
present invention, including mature forms, as measured in a particular biological assay,
with or without dose dependency. In the case where dose dependency does exist, it
need not be identical to that of the polypeptide, but rather substantially similar to the
dose-dependence in a given activity as compared to the polypeptide of the present
20 invention (i.e., the candidate polypeptide will exhibit greater activity or not more than
about 25-fold less and, preferably, not more than about tenfold less activity, and most
preferably, not more than about three-fold less activity relative to the polypeptide of the
present invention.)

25 Polynucleotides and Polypeptides of the Invention

FEATURES OF PROTEIN ENCODED BY GENE NO: 1

This gene is expressed primarily in testes tumor and to a lesser extent in fetal brain.

30 Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, cancer particularly of the testes, and defects of the central nervous system
such as seizure and neurodegenerative disorders. Similarly, polypeptides and antibodies
35 directed to these polypeptides are useful in providing immunological probes for
differential identification of the tissue(s) or cell type(s). For a number of disorders of
the above tissues or cells, particularly cancer of the testes and central nervous system,

expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., testes and other reproductive tissue, brain and other tissue of the nervous system, and blood cells, and spleen, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment/diagnosis of testicular cancer and treatment of central nervous system disorders since this gene is primarily expressed in the testes tumor and developing brain.

FEATURES OF PROTEIN ENCODED BY GENE NO: 2

This gene is expressed primarily in cancer tissues, such as breast cancer and Wilm's tumor, and to a lesser extent in fetal tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, and/or tumors, particularly, those found in the breast, and developmental abnormalities or disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the glandular tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., mammary tissue, and fetal tissue and, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 314 as residues: Pro-11 to Thr-18, Leu-43 to Pro-50, Gly-64 to Leu-72, and Leu-81 to Lys-86.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment/diagnosis of cancers and/or tumors, particularly, those found in the breast since expression is mainly in cancer/tumor tissues. May serve as therapeutic proteins for proliferation/differentiation of fetal tissues.

FEATURES OF PROTEIN ENCODED BY GENE NO: 3

This gene is expressed primarily in CD34 depleted buffy coat and to a lesser extent in spleen, chronic lymphocytic leukemia.

- 5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: blood disorders or leukemias, diseases of the immune system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for
- 10 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, and spleen, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or
- 15 cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment/diagnosis of blood disorders or
- 20 leukemias, diseases of the immune system since expression is in tissues related to immune function.

FEATURES OF PROTEIN ENCODED BY GENE NO: 4

This gene is expressed primarily in CD34 depleted buffy coat.

- 25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: blood disorders or lymphocytic diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification
- 30 of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual
- 35 having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment/diagnosis of blood disorders since expression is in tissues related to immune function.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 5

This gene is expressed primarily in CD34 depleted buffy coat.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: blood or immune
10 diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, and cancerous
15 and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 317 as residues:
20 Pro-13 to Lys-21.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment/diagnosis of blood disorders since expression is in tissues related to immune function.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 6

This gene is expressed primarily in CD34 depleted buffy coat.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: blood or immune
30 diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., and blood cells, and
35 cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level

in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 318 as residues: Lys-31 to Lys-39.

5 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment/diagnosis of blood diseases since it is expressed in tissues related to immune function.

FEATURES OF PROTEIN ENCODED BY GENE NO: 7

10 This gene is expressed primarily in CD34 depleted buffy coat and to a lesser extent in pineal gland.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: diseases of the immune system and brain associated diseases. Similarly, polypeptides and antibodies directed to
15 these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, and pineal gland, and cancerous and wounded tissues) or
20 bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment/diagnosis of blood disorders, immune diseases or brain associated diseases (specifically of the pineal gland) since expression is in tissues related to immune function.

FEATURES OF PROTEIN ENCODED BY GENE NO: 8

30 The translation product of this gene shares sequence homology with an organic cation transporter which is thought to be important in organic cation uptake in the kidney and liver. (See Accession No. 2343059.) Preferred polypeptide fragments comprise the amino acid sequence ITIAIQMICLVNXELYPTFVRNXGVMVCSSLCDIGGIITP FIVFRLREVWQALPLILFAVLGLLAAGVTLLLPETKGVLPETMKDAENLGRKAKPKENTIYLLK
35 VQTSEPSGT (SEQ ID NO: 615) or TMKDAENLGRKAKPKENT (SEQ ID NO: 616) as well as N-terminal and C-terminal deletions of these fragments. Also preferred are polynucleotide fragments encoding these polypeptide fragments.

This gene is expressed primarily in liver.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: hepatic and renal diseases where drug elimination/cation exchange (organic cation uptake) in the liver and kidney are problematic. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hepatic or renal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., kidney and liver, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 320 as residues: Asn-64 to Asn-74, and Gln-81 to Gly-87.

The tissue distribution and homology to organic cation transporter indicate that polynucleotides and polypeptides corresponding to this gene are useful as a polyspecific transporter that is important for drug elimination in the liver (and possibly kidney) since expression is found in the liver.

FEATURES OF PROTEIN ENCODED BY GENE NO: 9

This gene is expressed primarily in eosinophil induced with IL-5 and to a lesser extent in fetal liver and spleen. This gene also maps to chromosome 15, and therefore can be used in linkage analysis as a marker for chromosome 15.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: diseases of the immune system, particularly allergies or asthma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, liver, and spleen, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the

standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating/diagnosis of diseases involving eosinophil reactions since expression seems to be concentrated in eosinophils and other tissues involved in immunity such as the liver and spleen.

FEATURES OF PROTEIN ENCODED BY GENE NO: 10

This gene is expressed primarily in tissues of hematopoietic lineage and to a lesser extent in Hodgkins lymphoma. Any frame shifts in this sequence can easily be clarified using known molecular biology techniques.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, and immune deficiency or dysfunction. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., hematopoietic cells, lymphoid and reticuloendothelial tissues, and cancerous tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment/ diagnosis for lymphomas or immune dysfunction or as a therapeutic protein useful in immune modulation based on expression in anergic T-cells and lymphomas.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 11

This gene is expressed primarily in neutrophils and to a lesser extent in activated lymphoid cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the cell type present in a biological sample and for diagnosis of diseases and conditions: inflammation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders

of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells and lymphoid tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another
5 tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 323 as residues: Glu-40 to Lys-46.

The tissue distribution indicates that polynucleotides and polypeptides
10 corresponding to this gene are useful for modulation of an immune reaction or as a growth factor for the differentiation or proliferation of neutrophils for the treatment of neutropenia.

FEATURES OF PROTEIN ENCODED BY GENE NO: 12

15 This gene is expressed primarily in brain and to a lesser extent in activated T-cells. It is likely that the open reading frame containing the predicted signal peptide continues in the 5' direction. Preferred polypeptide fragments comprise the amino acid sequence PRVRNSPEDLGLSLTGDSCKL (SEQ ID NO:617).

Therefore, polynucleotides and polypeptides of the invention are useful as
20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: neurodegenerative disorders including ischemic shock, alzheimers and cognitive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a
25 number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, and brain, and other tissue of the nervous system and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from
30 an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 324 as residues: Ser-5 to Glu-14, Ile-21 to Pro-35, Ser-65 to Asp-81, Cys-89 to Val-96, Lys-136 to Ser-145, Ile-152 to Met-169, and Arg-189 to Lys-196.

35 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnostic/treatment for cancers of the given tissue or in the treatment of neurological disorders of the CNS.

FEATURES OF PROTEIN ENCODED BY GENE NO: 13

This gene was also recently cloned by other groups, naming this calcium-activated potassium channel gene, hKCa4. (See Accession No. AF033021, see also, Accession
5 No. 2584866.) This gene is mapped to human chromosome 19q13.2. A second signal sequence likely exists upstream from the predicted signal sequence as described in Table 1. Preferred polypeptide fragments comprise: QADDLQATVAALCVLRGGGPWAG SWLSPKTPGAMGGDLVLGLGALRRRKRL (SEQ NO: 618); or EQEKSLAGWALVLAXXGIGL MVLHAEMLWFGGCSAVNATGHLSDTLWLIPITFLTIGYGDVVPGTMWGKIVCLCTGVMGVCC
10 TALLVAVVARKLEFNKAEKHVHNFMMDIQYTKEMKESAARVLQEAWMFYKHTRRKESHAAR XHQRXLLAAINAFRQVRLKHKRLREQVNSMVDISKMHMILYDLQQNLSSSHRALEKQIDTLAG KLDALTELLSTALGPRQLPEPSQQSK (SEQ ID NO: 619), as well as N-terminal and C-terminal deletions. Also preferred are polynucleotide fragments encoding these polypeptide fragments.

15 This gene is expressed primarily in breast lymph node and T-cells, and to a lesser extent in placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: hematologic and
20 immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., lymphoid
25 tissue, blood cells and placenta, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a
30 sequence shown in SEQ ID NO. 325 as residues: Arg-13 to Lys-23.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment/diagnosis of hematologic and diseases involving immune modulation based on distribution in the lymph node and T-cells.

FEATURES OF PROTEIN ENCODED BY GENE NO: 14

This gene was recently cloned by another group, calling it PAPS synthase. (See Accession No. e1204135.) Preferred polypeptide fragments comprise the amino acid sequence YQAHVSRNKRQVVGTRGGFRGCTVWLTGLSGAGK (SEQ ID NO: 620).

5 Also preferred are the polynucleotide fragments encoding this polypeptide fragment.

It has been discovered that this gene is expressed primarily in benign prostate hyperplasia, Human Umbilical Vein Endothelial Cells and to a lesser extent in smooth muscle and Human endometrial stromal cells-treated with estradiol.

Therefore, polynucleotides and polypeptides of the invention are useful as
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: inflammation, ischemia, and restenosis, based on endothelial cell and smooth muscle cell expression, and prostate diseases such as benign prostate hyperplasia or prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing
15 immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the prostate or vessels of the circulatory system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., prostate, endothelial cells, smooth muscle, and endometrium, and cancerous and wounded tissues) or bodily
20 fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 326 as residues: Arg-21 to Asp-26, Lys-35 to Lys-44,
25 Glu-49 to Asn-58.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating/diagnosing diseases or conditions where the endothelial cell lining of the veins and arteries of underlying smooth muscle are involved.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 15

This gene is expressed primarily in human 6 week embryo and to a lesser extent in placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as
35 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: developmental anomalies or fetal deficiencies. Similarly, polypeptides and antibodies directed to these

polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly developmental in nature, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., embryonic
5 tissue, and placenta, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID
10 NO. 327 as residues Lys-50 to Glu-57.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for detection of developmental abnormalities.

FEATURES OF PROTEIN ENCODED BY GENE NO: 16

15 This gene is expressed primarily in kidney and amygdala and to a lesser extent in fetal tissues. This gene is mapped to chromosome 14, and therefore is useful in linkage analysis as a marker for chromosome 14.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) present in a biological sample and
20 for diagnosis of diseases and conditions: kidney diseases, neurological disorders and developmental abnormalities. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s). For a number of disorders of the above tissues, particularly of the renal system or developing fetal tissues, expression of this gene at significantly higher or
25 lower levels may be routinely detected in certain tissues and cell types (e.g., kidney, amygdala, and fetal tissues, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an
30 individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment or diagnosis of conditions affecting the brain, kidneys and fetal development.

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 17

This gene is expressed primarily in ovarian cancer.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: solid tumors similar to ovarian cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., ovarian and other reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 329 as residues Ser-51 to Val-56.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment of solid tumors of the reproductive system such as ovarian cancer.

FEATURES OF PROTEIN ENCODED BY GENE NO: 18

This gene is expressed primarily in brain medulloblastoma. Preferred polypeptide fragments comprise the amino acid sequence: IRHEQHPNFSLEMHSGSSLLLFLPQL ILILPVCAHLHEELNC (SEQ ID NO: 643) and SFFISEEKGHLLLQAERHPWVAGALVGVSGLTLTTCSGPTEKPKATKNYFLKRLLQEMHIRAN (SEQ ID NO: 644), as well as N-terminal and C-terminal deletions. Also preferred are polynucleotide fragments encoding these polypeptide fragments.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: tumors particularly of the CNS or. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the Central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating medulloblastoma or similar tumors.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 19

This gene is expressed primarily in adipocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: obesity. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the adipose tissues expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., adipocytes and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating obesity by regulating the function and number of adipocytes

FEATURES OF PROTEIN ENCODED BY GENE NO: 20

25 This gene is expressed primarily in B cell lymphoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, of the immune system with an emphasis on B cell lymphoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the tumors of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, and lymphoid tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,

30
35

the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of B cell derived tumors based on its expression in b cell lymphomas

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

This gene is expressed primarily in immune cells and to a lesser extent in fetal tissues

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: inflammatory diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., cells of the immune system, and fetal tissues, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:333 as residues Asp-10 to Pro-19, Ser-74 to Tyr-79, Glu-95 to Lys-110.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment of diseases involving alterations in T cell activity.

FEATURES OF PROTEIN ENCODED BY GENE NO: 22

It has been discovered that this gene is expressed primarily in ovarian tumor.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: tumors particularly of the ovary. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of tumors of the reproductive organs. expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., ovarian

and other reproductive tissue and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 334 as residues: Leu-22 to Gln-27.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of ovarian tumors as it has only been identified in ovarian tumors.

FEATURES OF PROTEIN ENCODED BY GENE NO: 23

It has been discovered that this gene is expressed primarily in fetal tissues and to a lesser extent in osteoclastoma cell line

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: osteoporosis or arthritis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., bone cells, and fetal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment of conditions of abnormal bone remodeling due to enhanced activity of osteoclasts. This may be useful as a specific marker for malignancies derived from osteoclasts or their precursors.

FEATURES OF PROTEIN ENCODED BY GENE NO: 24

The translation product of this gene shares sequence homology with a periplasmic ribonuclease which is thought to be important in degrading extracellular polynucleotides

It has been discovered that this gene is expressed primarily in serum treated smooth muscle cells

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: vascular disease such as restenosis. Similarly, polypeptides and antibodies directed to these polypeptides are
5 useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vasculature expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., smooth muscle, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or
10 spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 336 as residues: Gln-30 to Lys-36, and Pro-41 to Arg-48.

15 The tissue distribution and homology to ribonucleases indicate that polynucleotides and polypeptides corresponding to this gene are useful for treatment of pathological conditions of smooth muscle associated with bacterial or viral infiltration

FEATURES OF PROTEIN ENCODED BY GENE NO: 25

20 This gene is expressed primarily in Early Stage Human Brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: human brain development and related diseases. Similarly, polypeptides and antibodies directed to
25 these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the human brain development and related diseases, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, and
30 cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to this gene indicate that polynucleotides
35 and polypeptides corresponding to this gene are useful for diagnosis and treatment of diseases affecting human brain development and related diseases.

FEATURES OF PROTEIN ENCODED BY GENE NO: 26

It has been discovered that this gene is expressed primarily in human brain tissue.

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: human brain diseases and other diseases related to brain diseases, which may be caused by brain diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in
10 providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the human brain diseases, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum,
15 plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 The tissue distribution and homology to the gene indicate that polynucleotides
20 and polypeptides corresponding to this gene are useful for diagnosis and treatment of human brain diseases and other diseases related.

FEATURES OF PROTEIN ENCODED BY GENE NO: 27

It has been discovered that this gene is expressed primarily in Anergic T-cells.

25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: immune diseases, inflammatory diseases and diseases related to T lymph cells. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological
30 probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune diseases, inflammatory diseases and diseases related to T lymph cells, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, and cancerous and wounded tissues) or bodily fluids (e.g.,
35 serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to the gene indicate that polynucleotides and polypeptides corresponding to this gene are useful for immune diseases,
5 inflammatory diseases and diseases related to T lymph cells.

FEATURES OF PROTEIN ENCODED BY GENE NO: 28

The translation product of this gene shares sequence homology with *Shigella flexneri* positive transcriptional regulator CriR (criR) gene which is thought to be
10 important in regulation of gene expression.

This gene is expressed primarily in human synovial sarcoma and normal human brain tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
15 biological sample and for diagnosis of diseases and conditions: human brain diseases particularly sarcomas of the synovium. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the human brain and synovium and other related human
20 brain diseases, expression of this gene at significantly higher or lower levels may be routinely detected in certain (e.g., synovial tissue, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,
25 the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of human synovial sarcoma and other related human brain diseases.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 29

This gene is expressed in bone marrow, infant brain, fetal liver and spleen, prostate and to a lesser extent in pineal gland, adipose tissue, kidney, adrenal gland, umbilical vein endothelial cells, and T cells.

35 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: diseases related to bone marrow or

hematoplastic tissues, prostate, kidney, adrenal gland, and cardiovascular tissue or organs. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the diseases related to

5 hematoplastic tissues, immune system, prostate, kidney, adrenal gland, and cardiovascular tissue or organs, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., bone marrow, hematopoietic cells, pineal gland, adipose tissue, kidney, adrenal gland, endothelial cells, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g.,

10 serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to the gene indicate that polynucleotides

15 and polypeptides corresponding to this gene are useful for diagnosis and treatment of diseases related to hematoplastic tissues, immune system, prostate, kidney, adrenal gland, and cardiovascular tissue or organs.

FEATURES OF PROTEIN ENCODED BY GENE NO: 30

20 This gene is expressed primarily in meningea and to a lesser extent in breast and adult brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: Diseases of the

25 meningea and related brain diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the meningea and related brain diseases, expression of this gene at significantly higher or lower levels may be routinely detected in certain

30 tissues and cell types (e.g., meningea, mammary tissue, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the

35 disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of diseases of the meningea and related brain diseases.

5 **FEATURES OF PROTEIN ENCODED BY GENE NO: 31**

This gene is expressed in meningea, fetal spleen, osteoblast and to a lesser extent in activated T-cells, endometrial stromal cells, fetal lung, HL-60, thymus, testis and endothelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: meningeal disease, osteoporosis, immune diseases, and hematoplastic diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for identification of the tissue(s) or cell type(s). For a number of disorders of the
15 above tissues or cells, particularly of the meningeal diseases, osteoporosis, immune diseases, and hematoplastic diseases, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, endometrium, lung, thymus, testis, and endothelial cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal
20 fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to gene indicate that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of
25 meningeal, osteoporosis, immune diseases, hematoplastic diseases, testis diseases and lung diseases.

FEATURES OF PROTEIN ENCODED BY GENE NO: 32

This gene is expressed primarily in human thymus and to a much lesser extent
30 in infant brain, T-cells, smooth muscle, endothelial cells, bone marrow, human ovarian tumor and keratinocytes testes, osteoclastoma, breast, and tonsils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: Diseases involving the
35 thymus, particularly thymic cancer and diseases involving T-cell maturation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a

number of disorders of the above tissues or cells, particularly of the thymus, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., thymus, brain, and other tissue of the nervous system, blood cells, bone marrow, ovaries, and testes, and other reproductive tissue, mammary
5 tissue, tonsils, melanocytes and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10 The tissue distribution and homology to gene indicate that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of diseases of the thymus particularly thymic cancer and diseases involving T-cell maturation.

15 **FEATURES OF PROTEIN ENCODED BY GENE NO: 33**

 This gene is expressed primarily in human tonsils, and placenta, and to a lesser extent in adipocytes, melanocyte, and infant brain.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
20 biological sample and for diagnosis of diseases and conditions: inflammatory diseases, immune diseases, and obesity. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the inflammatory diseases, immune diseases, and obesity, expression of
25 this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., tonsils, placenta, adipocytes, melanocytes, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard
30 gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 The tissue distribution and homology to this gene indicate that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of diseases such as inflammation, immune diseases, and obesity.

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 34

This gene is expressed in activated T cells, and to a lesser extent in pituitary, testis, and breast lymph node.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: diseases relating to T cells. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the disorders of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., pituitary, testes and other reproductive tissue, mammary tissue, and lymphoid tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment of immune disorders.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 35

This gene is expressed primarily in infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the diseases relating to neurological disorders, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain, and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

35 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of neurological disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 36

This gene is expressed primarily in infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions: neurological disorders.
Similarly, polypeptides and antibodies directed to these polypeptides are useful in
providing immunological probes for differential identification of the tissue(s) or cell
10 type(s). For a number of disorders of the above tissues or cells, particularly of the
diseases relating to neurological disorders, expression of this gene at significantly
higher or lower levels may be routinely detected in certain tissues and cell types (e.g.,
brain and other tissue of the nervous system, and cancerous and wounded tissues) or
bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another
tissue or cell sample taken from an individual having such a disorder, relative to the
15 standard gene expression level, i.e., the expression level in healthy tissue or bodily
fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides
corresponding to this gene are useful for diagnosis and treatment of neurological
disorders.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 37

This gene is expressed primarily in human ovary.

Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
25 biological sample and for diagnosis of diseases and conditions: ovarian cancer.
Similarly, polypeptides and antibodies directed to these polypeptides are useful in
providing immunological probes for differential identification of the tissue(s) or cell
type(s). For a number of disorders of the above tissues or cells, particularly of the
ovarian disorders such as those involving germ cells, ovarian follicles, stromal cells,
30 expression of this gene at significantly higher or lower levels may be routinely detected
in certain tissues and cell types (e.g., ovary and other reproductive tissue, and
cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial
fluid or spinal fluid) or another tissue or cell sample taken from an individual having
such a disorder, relative to the standard gene expression level, i.e., the expression level
35 in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides
corresponding to this gene are useful for diagnosis and treatment of ovarioathy.

FEATURES OF PROTEIN ENCODED BY GENE NO: 38

This gene is expressed primarily in lymph node breast cancer.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions: breast cancer. Similarly,
polypeptides and antibodies directed to these polypeptides are useful in providing
immunological probes for differential identification of the tissue(s) or cell type(s). For a
10 number of disorders of the above tissues or cells, particularly of the breast cancer,
expression of this gene at significantly higher or lower levels may be routinely detected
in certain tissues and cell types (e.g., mammary tissue and lymphoid tissue, and
cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial
fluid or spinal fluid) or another tissue or cell sample taken from an individual having
15 such a disorder, relative to the standard gene expression level, i.e., the expression level
in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides
corresponding to this gene are useful for used as a diagnostic marker for breast cancer.

FEATURES OF PROTEIN ENCODED BY GENE NO: 39

20 This gene is expressed primarily in brain and to a lesser extent in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions: neuronal disorders such
as trauma, brain degeneration, and brain tumor. Similarly, polypeptides and antibodies
25 directed to these polypeptides are useful in providing immunological probes for
differential identification of the tissue(s) or cell type(s). For a number of disorders of
the above tissues or cells, particularly of the brain, expression of this gene at
significantly higher or lower levels may be routinely detected in certain tissues and cell
types (e.g., brain and other tissue of the nervous system, and cancerous and wounded
30 tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or
another tissue or cell sample taken from an individual having such a disorder, relative to
the standard gene expression level, i.e., the expression level in healthy tissue or bodily
fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides
35 corresponding to this gene are useful for diagnosis and therapeutic treatment of
neuronal disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 40

5 This gene is expressed in early stage human embryo, adrenal gland tumor, and immune tissues such as fetal liver, fetal spleen, T-cell, and myeloid progenitor cell line and to a lesser extent in ovary, colon cancer, and a few other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: tumorigenesis including
10 adrenal gland tumor, colon cancer and various other tumors, developmental and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cancer tissues, early stage human tissues, and immune system,
15 expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, spleen, blood cells, bone marrow, ovary and other reproductive tissue, and colon, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard
20 gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and therapeutic treatment of immune and developmental disorders, and tumorigenesis.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 41

This gene is expressed primarily in fetal lung, endothelial cells, liver, thymus and a few other immune tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as
30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: immune disorders such as immune deficiency and autoimmune diseases, pulmonary diseases, liver diseases, and tumor matasis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification
35 of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal lung, liver, endothelial cells, and immune tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain

tissues and cell types (e.g., lung, endothelial cells, liver, thymus, and other tissue of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,
5 the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis of immune disorders and pulmonary and hepatic diseases. Its promoter may also be used for immune system and lung-
10 specific gene therapies. The expression of this gene in endothelial cells indicates that it may also involve in angiogenesis which therefore may play role in tumor matasis.

FEATURES OF PROTEIN ENCODED BY GENE NO: 42

This gene is expressed primarily in liver, thyroid, parathyroid and to a lesser
15 extent in fetal lung, stomach and early embryos.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: metabolic regulation, obesity, hepatic failure, hepatocellular tumors or thyroiditis and thyroid tumors. Similarly,
20 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive/endocrine system expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, thyroid, parathyroid, lung,
25 stomach, and embryonic tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30 The tissue distribution and the extracellular locations indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of digestive/endocrine disorders, including metabolic regulation, hepatic failure, malabsorption, gastritis and neoplasms.

FEATURES OF PROTEIN ENCODED BY GENE NO: 43

This gene is expressed primarily in Schizophrenic adult brain, pituitary, front cortex, hypothalamus and to a lesser extent in retina, adipose and stomach cancer and placenta.

- 5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: schizophrenia and other neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification
- 10 of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nerve system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., retinal tissue, adipose, stomach, and placenta, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or
- 15 cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful in treatment/detection of disorders in the nerve
- 20 system, including schizophrenia, neurodegeneration, and neoplasia. Additionally, a secreted protein in brain may serve as an endocrine.

FEATURES OF PROTEIN ENCODED BY GENE NO: 44

- The translation product of this gene shares sequence homology with GTP
- 25 binding proteins which are thought to be important in signal transduction and protein transport.

This gene is expressed primarily in umbilical vein and microvascular endothelial cells, GM-CSF treated macrophage, anergic T cells, osteoblast, osteoclast, CD34+ cells and to a lesser extent in gall bladder.

- 30 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: bone formation and growth, osteonecrosis, osteoporosis, angiogenesis and/or hematopoiesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing
- 35 immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal and hematopoiesis systems, expression of this gene at significantly higher or lower levels

may be routinely detected in certain tissues and cell types (e.g., endothelial cells, blood cells, bone, and gall bladder, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene
5 expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to GTP binding proteins indicate that polynucleotides and polypeptides corresponding to this gene are useful for treatment/detection of bone formation and growth, osteonecrosis, osteoporosis, and/or
10 hematopoiesis because its involvement in the growth signaling or angiogenesis.

FEATURES OF PROTEIN ENCODED BY GENE NO: 45

The translation product of this gene shares sequence homology with signal sequence receptor gamma subunit which is thought to be important in protein
15 translocation on endoplasmic reticulum.

This gene is expressed primarily in adrenal gland, salivary gland, prostate, and to a lesser extent in endothelial cells and smooth muscle.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
20 biological sample and for diagnosis of diseases and conditions: protein secretion. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the secretory organs, expression of this gene at significantly higher or lower levels may be
25 routinely detected in certain tissues and cell types (e.g., adrenal gland, salivary gland, prostate, endothelial cells, and smooth muscle, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily
30 fluid from an individual not having the disorder.

The tissue distribution and homology to SSR gamma subunit indicate that polynucleotides and polypeptides corresponding to this gene are useful for endocrine disorders, prostate cancer, xerostomia or sialorrhea.

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 46

This gene is expressed primarily in osteoclastoma cells and to a lesser extent in melanocyte, amygdala, brain, and stomach.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: ossification, osteoporosis, fracture, osteonecrosis, osteosarcoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., melanocytes, amygdala, brain and other tissue of the nervous system, and stomach, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful in intervention of ossification, osteoporosis, fracture, osteonecrosis and osteosarcoma.

FEATURES OF PROTEIN ENCODED BY GENE NO: 48

The translation product of this gene shares sequence homology with proline rich proteins which is thought to be important in protein-protein interaction.

This gene is expressed primarily in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: neurological and psychological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nerve system and endocrine system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to proline-rich proteins indicate that polynucleotides and polypeptides corresponding to this gene are useful in intervention

and detection of neurological diseases, including trauma, neoplasia, degenerative or metabolic conditions in the central nerve system. Additionally, the gene product may be a secreted by the brain as an endocrine.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 49

The translation product of this gene shares sequence homology with the AOCB gene from *Aspergillus nidulans* which is important in asexual development.

This gene is expressed primarily in infant brain and to a lesser extent in the developing embryo, trachea tumors, B-cell lymphoma and synovial sarcoma.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: neurodegenerative diseases, leukemia and sarcoma's. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential
15 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., embryonic tissue, blood cells, trachea, and synovial tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or
20 spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in infant brain and sarcoma's and homology to a gene involved in a key step of eukaryotic development (fungal spore formation) indicates
25 that the protein product of this clone could play a role in neurological diseases such as schizophrenia, particularly in infants. The existence of the gene in a B-cell lymphoma indicates the gene may be used in the treatment and detection of leukemia.

FEATURES OF PROTEIN ENCODED BY GENE NO: 50

30 This gene is expressed primarily in fetal lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: pulmonary disorders including lung cancer. Similarly, polypeptides and antibodies directed to these
35 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the pulmonary system, expression of this gene at significantly higher or

lower levels may be routinely detected in certain tissues and cell types (e.g., lung, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of this gene only in fetal lung indicates that it plays a key role in development of the pulmonary system. This would suggest that misregulation of the expression of this protein product in the adult could lead to lymphoma or sarcoma formation, particularly in the lung. It may also be involved in predisposition to certain pulmonary defects such as pulmonary edema and embolism, bronchitis and cystic fibrosis.

FEATURES OF PROTEIN ENCODED BY GENE NO: 51

This gene is expressed primarily in hematopoietic cell types and fetal cells and to a lesser extent in all tissue types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: defects in the immune system and hematopoiesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., hematopoietic cells, and fetal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of this gene predominantly in hematopoietic cells and in the developing embryo indicates that polynucleotides and polypeptides corresponding to this gene are useful for detection and treatment of lymphomas and disease states affecting the immune system or hematopoiesis disorders such as leukemia, AIDS, arthritis and asthma..

FEATURES OF PROTEIN ENCODED BY GENE NO: 52

This gene is expressed primarily in prostate and to a lesser extent in fetal spleen, fetal liver, infant brain and T cell leukemias.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: prostate disorders, prostate cancer, leukemia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, and/or prostate gland expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., thymus, spleen, liver, brain and other tissue of the nervous system, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of this gene in prostate indicates that polynucleotides and polypeptides corresponding to this gene are useful for detection or treatment of prostate disorders or prostate cancer. Its distribution in fetal liver and fetal spleen indicates it may play a role in the immune system and its misregulation could lead to immune disorders such as leukemia, arthritis and asthma.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 53

The translation product of this gene shares sequence homology with dynein.

This gene is expressed primarily in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: neuro-degenerative diseases of the brain. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly neuro-degenerative diseases expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The predominant tissue distribution in the brain and homology to dynein, a microtubule motor protein involved in the positioning of cellular organelles and molecules indicates that polynucleotides and polypeptides corresponding to this gene are useful for detection/treatment of neurodegenerative diseases, such as Alzheimers,
5 Huntigtons, Parkinsons diseases and shizophrenia.

FEATURES OF PROTEIN ENCODED BY GENE NO: 54

The translation product of this gene shares sequence homology with ubiquitin-conjugation protein, an enzyme which is thought to be important in the processing of
10 the Huntingtons Disease causing gene.

This gene is expressed primarily in brain and to a lesser extent in activated macrophages.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
15 biological sample and for diagnosis of diseases and conditions: neurodegenerative disease states including Huntington's disease. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of brain tissues. For a number of disorders of the above tissues or cells, particularly of the neurological systems expression of this gene at
20 significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level
25 in healthy tissue or bodily fluid from an individual not having the disorder.

The predominant tissue distribution of this gene in the brain and its homology to a Huntington interacting protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for the regulation of the expression of the Huntington disease gene and other neurodegenerative diseases including
30 spinocerebellar ataxia types I and III, dentatorubropallidoluysian and spinal bulbar muscular atrophy. In addition, the existence of elevated levels of free ubiquitin pools in Alzheimer's disease, Parkinson's disease and amyotrophic lateral sclerosis indicates that the ubiquitin pathway of protein degradation plays a role in these disease states. Thus, considering the gene described here is homologous to a ubiquitin-conjugation
35 protein it may play a general role in neurodegenerative conditions.

FEATURES OF PROTEIN ENCODED BY GENE NO: 56

This gene is expressed primarily in T-cells (anergic T-cells, resting T-Cells, apoptotic T-cells) and lymph node (breast), as well as brain (hypothalamus, hippocampus, pituitary, infant brain, early-stage brain).

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: immune (e.g. immunodeficiencies, autoimmunities, inflammation, leukemias & lymphomas) and neurological (e.g. Alzheimer's disease, dementia, schizophrenia) disorders. Similarly,
10 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous, hematopoietic and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood
15 cells, lymphoid tissue, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful in the intervention or detection of pathologies associated with the hematopoietic and immune systems, such as anemias (leukemias). In addition, the expression in brain (including fetal) might suggest a role in developmental brain defects, neuro-degenerative diseases or behavioral abnormalities
25 (e.g. schizophrenia, Alzheimer's, dementia, depression, etc.).

FEATURES OF PROTEIN ENCODED BY GENE NO: 57

This gene is expressed primarily in lung, and to a lesser extent in a variety of other hematological cell types (e.g. Raji cells, bone marrow cell line, activated
30 monocytes).

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: pulmonary and/or hematological disfunction. Similarly, polypeptides and antibodies directed to these
35 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vasculo-pulmonary and hematopoietic systems, expression of this

gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., lung and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the
5 standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful in the intervention and detection of pathologies associated with the vasculo-pulmonary system. In addition the expression of this gene
10 in a variety of leukocytic cell types and a bone marrow cell line might suggest a role in hematopoietic and immune system disorders, such as leukemias & lymphomas, inflammation, immunodeficiencies and autoimmunities.

FEATURES OF PROTEIN ENCODED BY GENE NO: 58

15 The translation product of this gene shares sequence homology with adenylate kinase isozyme 3 (gil163528 GTP:AMP phosphotransferase (EC 2.7.4.10) [Bos taurus]), which is thought to be important in catalyzing the phosphorylation of AMP to ADP in the presence of ATP or inorganic triphosphate.

This gene is expressed primarily in fetal liver, heart and placenta, and to a lesser
20 extent in many other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: hepatic, cardiovascular or reproductive disorders. Similarly, polypeptides and antibodies directed to these
25 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hepatic, cardiovascular and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, heart, and placenta, and cancerous and wounded tissues) or
30 bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides
35 corresponding to this gene are useful for the treatment and diagnosis of conditions related to hepatic function and pathogenesis, in particular, those dealing with liver development and the differentiation of hepatocyte progenitor cells.

FEATURES OF PROTEIN ENCODED BY GENE NO: 59

This gene is expressed primarily in CD34 positive cells (Cord Blood).

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions: hematopoietic
differentiation and immune disorders. Similarly, polypeptides and antibodies directed to
these polypeptides are useful in providing immunological probes for differential
identification of the tissue(s) or cell type(s). For a number of disorders of the above
10 tissues or cells, particularly of hematopoietic and immune systems, expression of this
gene at significantly higher or lower levels may be routinely detected in certain tissues
and cell types (e.g., hematopoietic cells, and blood cells, and cancerous and wounded
tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or
another tissue or cell sample taken from an individual having such a disorder, relative to
15 the standard gene expression level, i.e., the expression level in healthy tissue or bodily
fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides
corresponding to this gene are useful in the detection and treatment of conditions
associated with CD34-positive cells, and therefore as a marker for cell differentiation in
20 hematopoiesis, as well as immunological disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 60

The translation product of the predicted open reading frame of this contig has
sequence identity to the murine gene designated Insulin-Like Growth Factor-Binding
25 Protein (IGFBP)-1 as described by Lee and colleagues (Hepatology 19 (3), 656-665
(1994)).

This gene is expressed exclusively in hemangiopericytoma.

Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
30 biological sample and for diagnosis of hemangiopericytoma and other pericyte or
endothelial cell proliferative disorders. Similarly, polypeptides and antibodies directed
to these polypeptides are useful in providing immunological probes for differential
identification of the tissue(s) or cell type(s). For a number of disorders of the above
tissues or cells, particularly of the circulatory and immune systems, expression of this
35 gene at significantly higher or lower levels may routinely be detected in certain tissues
and cell types (e.g., pericyte or endothelial cells, and liver, and cancerous and wounded
tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or

another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Polynucleotides and polypeptides corresponding to this gene are useful as cell growth regulators since IGFBP-1-like molecules function as modulators of insulin-like growth factor activity. In addition, since IGFBP-1 is expressed at high levels following hepatectomy and during fetal liver development, polynucleotides of the present invention may also be used for the diagnosis of developmental disorders. Further, polypeptides of the present invention may be used therapeutically to treat developmental liver disorders as well as to regulate hepatocyte and supporting cell growth following hepatectomy or to treat liver disorders.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of hemangiopericytoma and liver disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 61

This gene is expressed primarily in schizophrenic frontal cortex.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: nervous system and cognitive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the frontal cortex and CNS expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, treatment and diagnosis of frontal cortex, neuro-degenerative and CNS disorders

FEATURES OF PROTEIN ENCODED BY GENE NO: 62

This gene is expressed primarily in human adrenal gland tumor, and to a lesser extent in human kidney, medulla and adult pulmonary tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: metabolic, endocrine disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine and nervous system disorders and neoplasia, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., adrenal gland, kidney, brain and other tissue of the nervous system, pulmonary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, treatment and diagnosis of neurological and endocrine disorders including neoplasia.

FEATURES OF PROTEIN ENCODED BY GENE NO: 63

This gene is expressed primarily in human adipocytes, and to a lesser extent in spleen, 12-week old human, and testes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: immune, metabolic and growth disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., adipocytes, spleen, and testes and other reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, diagnosis and treatment of immune, developmental and metabolic disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 64

One translated product of this clone is homologous to the mouse zinc finger protein PZF. (See Accession No. 453376; see also Gene 152 (2), 233-238 (1995).) Preferred polypeptide fragments correspond to the highly conserved domains shared between mouse and man. For example, preferred polypeptide fragments comprise the amino acid sequence: LQCEICGFTCRQKASLNWHMKKHDADSFYQFSCNICGKKFEKKDSVVAHKAKSH PEV (SEQ ID NO: 621); ITSTDILGTNPESLTQPSD (SEQ ID NO: 622); NSTSGECLLLEAEGM SKSY (SEQ ID NO: 623); CSGTERVSLMADGKIFVSGSGSGGTEGLVMNSDILGATTEVLIEDSD SAGP (SEQ ID NO: 624); IQYVRCMEGCGTVLAHPRYLQHHLKQHHIKYQHLLKKKYVCPHPSCGRFL RLQKQLLRHAKHHT (SEQ ID NO: 625); DQRDYICEYCARAFKSSHNLAVHRMIHTGEK (SEQ ID NO: 626); RSSRTSVSRHRDTENTRSSRSKTGSLQLICKSEPNTDQLDY (SEQ ID NO: 627); PFKDDPRDETYKPHLERETPKPRRKSG (SEQ ID NO: 630); QYVRCMEGCGTVLAHPRYLQ HHIKYQHLLKKKYVCPHPSCGRFLRLQKQLLRHAKHHTD (SEQ ID NO: 629); or residues 151-182 of QRDYICEYCARAFKSSHNLAVHRMIHTGEKHY (SEQ ID NO: 628). Also preferred are polynucleotide fragments encoding these polypeptide fragments.

This gene is expressed primarily in Rhabdomyosarcoma, melanocyte and colon cancer tissue and to a lesser extent in smooth muscle, pancreatic tumor, and apoptotic T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to,. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hemopoetic, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., striated muscle, melanocytes, colon, smooth muscle, pancreas, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, diagnosis and treatment of cancer and hemopoetic disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 65

This gene is expressed primarily in human adipose and salivary gland tissue and to a lesser extent in human bone marrow and fetal kidney.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: metabolic and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the metabolic and hemopoetic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., adipose, salivary gland, bone marrow, and kidney, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, diagnosis of metabolic and immune disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 66

This translated product of this gene was recently identified as oxytocinase splice variant 1. (See Accession Nos. 2209276 and d1010078.) Preferred polypeptide fragments comprise the amino acid sequence: EMFDSL SYFKGSSLLMLKTYLSEDVFQHAVVLYLHN HSYASIQSDDLWDSFNEVTNQTL DVKRMMKTWTLQKGFPLVTVQKKGKELFIQQRFFLNMK PEIQPSDTRYM (SEQ ID NO: 631). Also preferred are polynucleotide fragments encoding this polypeptide fragment.

FEATURES OF PROTEIN ENCODED BY GENE NO: 67

This gene is expressed primarily in hemopoetic cells, particularly apoptotic T-cells, and to lesser extent in primary dendritic cells and adipose tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of apoptotic T-cells, primary dendritic cells, and adipose tissue present in a biological sample and for diagnosis of diseases and conditions: hemopoetic diseases including cancer and general immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell

type(s). For a number of disorders of the above tissues or cells, particularly of the oral and intestinal mucosa as well as hemopoetic and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., hematopoietic cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment of diseases of the immune system, including cancer, hemopoetic and infectious diseases.

FEATURES OF PROTEIN ENCODED BY GENE NO: 68

This gene is expressed primarily in kidney cortex and to a lesser extent in infant brain, heart, uterus, and blood.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of kidney tissue present in a biological sample and for diagnosis of diseases and conditions: soft tissue cancer, inflammation, kidney fibrosis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and endocrines systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., kidney, brain, and other nervous tissue, heart, uterus, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of cancer and fibroses.

FEATURES OF PROTEIN ENCODED BY GENE NO: 69

The translation product of this gene shares strong sequence homology with vertebrate and invertebrate protein tyrosine phosphatases.

This gene is expressed primarily in endometrial tumors, melanocytes, myeloid progenitors and to a lesser extent in infant brain, adipocytes, and several hematopoietic stem cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of transformed hematopoietic and epithelial cells present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, of skin and endometrium, leukemia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and hemopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., endometrium, melanocytes, bone marrow, adipocytes, hematopoietic cells, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and sequence similarity with tyrosine phosphatases indicate that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of cancer and hematopoietic disorders.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 70

This gene is expressed primarily in osteoclastoma, breast, and infant brain and to a lesser extent in various fetal and transformed bone, ovarian, and neuronal cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: degenerative conditions of the brain and skeleton. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and skeletal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., bone, mammary tissue, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of degenerative, neurological and skeletal disorders.

5 **FEATURES OF PROTEIN ENCODED BY GENE NO: 71**

This gene was originally cloned from tumor cell lines. Recently another group has also cloned this gene, calling it the human malignant melanoma metastasis-suppressor (KiSS-1) gene. (See Accession No. U43527.) Preferred polypeptide fragments comprise the amino acid sequence: LEKVASVGNSRPTGQQLESGLLA (SEQ ID NO: 632); VHREEASCYCQAEPGDL (SEQ ID NO: 633); RPALRQAGGGTREPRQKRWAGL (SEQ ID NO: 634); and AVNFRPQRSQSM (SEQ ID NO: 635). Any frame shifts can easily be resolved using known molecular biology techniques.

This gene is expressed primarily in many types of carcinomas and to a lesser extent in many normal organs.

15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissues(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer particularly melanomas, and other hyperproliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in
20 providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of transformed organ tissue, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or
25 another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. As a tumor suppressor gene, increase amounts of the polypeptide can be used to treat patients having a particular cancer.

30 The tissue distribution indicates that this gene and the translated product is useful for diagnosing and study of cancer.

FEATURES OF PROTEIN ENCODED BY GENE NO: 72

This gene is expressed primarily in striatum and to a lesser extent in adipocytes and hemangiopericytoma.

35 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of striatal cells present in a biological sample and for diagnosis of diseases and conditions: neurological, fat and lysosomal storage

diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., striatal tissue, adipocytes, and vascular tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis, study and treatment of neurodegenerative and growth disorders.

15 **FEATURES OF PROTEIN ENCODED BY GENE NO: 73**

This gene is expressed primarily in bone marrow stromal cells and to a lesser extent in smooth muscle, testes, endothelium, and brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of bone marrow present in a biological sample and for diagnosis of diseases and conditions: connective tissue and hematopoietic diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., bone marrow, stromal cells, smooth muscle, testes and other reproductive tissue, endothelium, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, diagnosis, and treatment of connective tissue and blood diseases.

FEATURES OF PROTEIN ENCODED BY GENE NO: 74

This gene is expressed primarily in brain, fetal liver and lung and to a lesser extent in retina, spinal chord, activated T-cells and endothelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of brain and regenerating liver present in a biological sample and for diagnosis of diseases and conditions: CNS and spinal chord injuries, immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,
10 particularly of the nervous and immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, liver, pulmonary tissue, blood cells, and endothelial cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from
15 an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of hematopoietic and
20 neurological conditions.

FEATURES OF PROTEIN ENCODED BY GENE NO: 75

The translation product of this gene shares sequence homology with GTP binding proteins (intracellular).

25 This gene is expressed primarily in bone marrow, brain, and melanocytes and to a lesser extent in various endocrine and hematopoietic tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: hematopoietic and
30 nervous system conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and immune, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., bone
35 marrow, melanocytes, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder,

relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to nucleotide binding factors indicate that polynucleotides and polypeptides corresponding to this gene are useful for study,
5 diagnosis, and treatment of brain degenerative, skin and blood diseases.

FEATURES OF PROTEIN ENCODED BY GENE NO: 76

This gene is expressed primarily in activated T-cells and to a lesser extent in retina, brain, and fetal bone.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of activated T-cells and developing brain present in a biological sample and for diagnosis of diseases and conditions: immune deficiencies and skeletal and neuronal growth disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes
15 for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous, immune, and skeletomuscular systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, brain and other tissue of the nervous system, retinal tissue, and bone, and cancerous and wounded tissues) or
20 bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides
25 corresponding to this gene are useful for diagnosis, study and treatment of cancer, urogenital, and brain degenerative diseases.

FEATURES OF PROTEIN ENCODED BY GENE NO: 77

This gene is expressed primarily in fetal liver, activated monocytes, osteoblasts
30 and to a lesser extent in synovial, brain, and lymphoid tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of myeloid and lymphoid present in a biological sample and for diagnosis of diseases and conditions: inflammation, immune deficiencies, cancer. Similarly, polypeptides and antibodies directed to these
35 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and skeleton, expression of this gene at significantly

higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, blood cells, bone, synovial tissue, brain and other tissue of the nervous system, and lymphoid tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample
5 taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, diagnosis, and treatment of lymphoid
10 and mesenchymal cancers and nervous system diseases.

FEATURES OF PROTEIN ENCODED BY GENE NO: 78

The translation product of this gene shares sequence homology with polymerase polyprotein precursor which is thought to be important in DNA repair and replication

15 This gene is expressed primarily in infant brain and to a lesser extent in tumors and tumor cell lines

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
20 not limited to, especially of the neural system and developing organs. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neural system expression of this gene at significantly higher or lower levels may be routinely detected
25 in certain (e.g., brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30 The tissue distribution and homology to polymerase polyprotein precursor indicate that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of cancers especially of the neural system and developing organs

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 79

This gene is expressed primarily in muscle and endothelial cells and to a lesser extent in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: vascular diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain (e.g., muscle, endothelial cells, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the vascular and neural system including cardiovascular and endothelial.

FEATURES OF PROTEIN ENCODED BY GENE NO: 80

This gene is expressed primarily in placenta and to a lesser extent in fetal liver. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: developmental disorders and disorder of the haemopoietic system, fetal liver and placenta. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of developmental disorders and disorder of the haemopoietic system, fetal liver and placenta, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., placenta and liver, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of developmental disorders and disorders of the haemopoietic system, fetal liver and placenta.

FEATURES OF PROTEIN ENCODED BY GENE NO: 81

This gene is expressed primarily in bone marrow, placenta and tissues and organs of the hematopoietic system.

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: disorders of the bone and haemopoietic system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification
10 of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, bone and hematopoietic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., bone marrow, placenta, and hematopoietic cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal
15 fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of disorders of the
20 immune, bone and hematopoietic system

FEATURES OF PROTEIN ENCODED BY GENE NO: 82

 The translation product of this gene shares sequence homology with secretory carrier membrane protein which is thought to be important in protein transport and
25 export. Any frame shifts in coding sequence can be easily resolved using standard molecular biology techniques. Another group recently cloned this gene, calling it SCAMP. (See Accession No. 2232243.)

 This gene is expressed primarily in prostate, breast and spleen, and to a lesser extent in several other tissues and organs.

30 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: disorders of the breast prostate and spleen. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification
35 of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly disorders of the breast prostate and spleen, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell

types (e.g., prostate, mammary tissue, and spleen, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to secretory carrier membrane protein indicate that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of disorders of the breast, prostate and spleen.

10 **FEATURES OF PROTEIN ENCODED BY GENE NO: 83**

This gene is expressed primarily in developing organs and tissue like placenta and infant brain and to a lesser extent in developed organs and tissue like cerebellum and heart.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: neurological diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neural system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., placenta, heart, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of diseases of the neural system including neurological disorders and cancer.

30 **FEATURES OF PROTEIN ENCODED BY GENE NO: 84**

The translation product of this gene shares sequence homology with ATPase 6 in *Trypanosoma brucei* which is thought to be important in metabolism.

This gene is expressed primarily in tumor and fetal tissues and to a lesser extent in melanocytes, kidney cortex, monocytes and ovary.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions: metabolism disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., fetal tissues, melanocytes, kidney, blood cells, ovary and other tissue of the reproductive system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to ATPase indicate that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of metabolism disorders, especially in fetal and tumor tissue growth.

FEATURES OF PROTEIN ENCODED BY GENE NO: 85

The translation product of this gene shares sequence homology with the immunoglobulin superfamily of proteins which are known to be important in immune response and immunity.

This gene is expressed primarily in stromal cells, colon cancer, lung, amygdala, melanocyte and to a lesser extent in a variety of other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: defects of stromal cell development and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the stromal cells, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., stromal cells, colon, lung, amygdala, and melanocytes, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to immunoglobulin indicate that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of immune system disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 86

The translation product of this gene shares sequence homology with transcription initiation factor eIF-4 gamma which is thought to be important in gene transcription.

This gene is expressed primarily in tumor tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: tumorigenesis.

Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly in tumor tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., endometrium and lung, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to transcription initiation factor eIF-4 gamma indicate that polynucleotides and polypeptides corresponding to this gene are useful for gene regulation in tumorigenesis.

FEATURES OF PROTEIN ENCODED BY GENE NO: 87

The translation product of this gene shares sequence homology at low level in prolines with secreted basic proline-rich peptide II-2 which is thought to be important in protein structure or inhibiting hydroxyapatite formation in vitro.

This gene is expressed primarily in endometrial tumor and fetal lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: endometrial tumors. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the muscular/skeletal and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., endometrium, and lung, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample

taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to secreted basic proline-rich peptide II-2 indicate that polynucleotides and polypeptides corresponding to this gene are useful for inhibiting hydroxyapatite formation or establishing cell/tissue structure.

FEATURES OF PROTEIN ENCODED BY GENE NO: 88

This gene is expressed primarily in: amniotic cells induced with TNF in culture; and to a lesser extent in colon tissue from a patient with Crohn's Disease; parathyroid tumor; activated T-cells; cells of the human Caco-2 cell line; adenocarcinoma; colon; corpus colosum; fetal kidney; pancreas tumor; fetal brain; early stage brain, and anergic T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: tumors. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system; e.g., tumors, expression of this gene at significantly higher or lower levels may be routinely detected in certain (e.g., amniotic cells, colon, kidney, pancreas, parathyroid, brain and other tissue of the nervous system, blood cells, hematopoietic cells, liver, spleen, bone, testes and other reproductive tissue, brain and other tissue of the nervous system, and epithelial cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein product of this clone is useful for modulating tumorigenesis and other immune system conditions such as disorders in immune response.

FEATURES OF PROTEIN ENCODED BY GENE NO: 89

This gene is expressed primarily in fetal liver/spleen and hematopoietic cells and to a lesser extent in brain, osteosarcoma, and testis tumor.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions: leukemia and hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., hematopoietic cells, liver, spleen, bone, testes, and other reproductive tissue, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of hematopoietic and immune disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 90

The translation product of this gene shares weak sequence homology with mouse Gcap1 protein which is developmentally regulated in brain.

This gene is expressed primarily in infant and adult brain and fetal liver/spleen and to a lesser extent in smooth muscle, T cells, and a variety of other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: neurological or hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous, hematopoietic, immune, and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, blood cells, liver, spleen, and smooth muscle, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and its homology to Gcap1 protein indicate that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders in neuronal, hematopoietic, immune, and endocrine systems.

5 **FEATURES OF PROTEIN ENCODED BY GENE NO: 91**

This gene is expressed primarily in brain and hematopoietic cells and to a lesser extent in tumor tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: disorder in nervous, hematopoietic, immune systems and tumorigenesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous, hematopoietic, immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein product of this clone is useful for diagnosis and treatment of disorders in the nervous, hematopoietic, and immune systems.

25 **FEATURES OF PROTEIN ENCODED BY GENE NO: 92**

The translation product of this gene shares sequence homology with neuroendocrine-specific protein A which is thought to be important in neurologic systems.

30 This gene is expressed primarily in brain tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: neural disorders and degeneration disease. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central or peripheral nervous systems, expression of this gene at

significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., hematopoietic cells, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having
5 such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to neuroendocrine-specific protein A indicate that polynucleotides and polypeptides corresponding to this gene are useful for treatment or diagnosis of neural disorders and degeneration disease.

10

FEATURES OF PROTEIN ENCODED BY GENE NO: 93

The translation product of this gene shares sequence homology with collagen-like protein and prolin-rich protein which are thought to be important in connective tissue function and tissue structure.

15 This gene is expressed primarily in fetal liver/spleen and brain tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: neuronal or hematopoietic disorders. Similarly, polypeptides and antibodies directed to these
20 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, spleen, and brain and other tissue of the nervous system, and
25 cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to collagen-like protein and proline-rich
30 proteins indicate that polynucleotides and polypeptides corresponding to this gene are useful for supporting brain and hematopoietic tissue function and diagnosis and treatment of disorders in these functions.

FEATURES OF PROTEIN ENCODED BY GENE NO: 94

35 This gene is expressed primarily in embryonic tissues and tumor tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to,. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system (e.g., tumors), expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., embryonic tissue and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of cancer.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 95

This gene is expressed primarily in brain tumor, placenta, and melanoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: brain tumor or melanoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain or melanocytes, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, placenta, and melanocytes, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the translation product of this gene is useful in the diagnosis and treatment of brain tumors and melanoma.

FEATURES OF PROTEIN ENCODED BY GENE NO: 96

The translation product of this gene shares sequence homology with a yeast membrane protein, SUR4, which encodes for APA1 that acts on a glucose-signaling pathway that controls the expression of several genes that are transcriptionally regulated by glucose.

This gene is expressed primarily in fetal liver, and to a lesser extent in placenta and breast tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: defects of fetal liver or defects of glucose-regulated ATPase activities in tissues. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal immune/hematopoietic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, placenta, and mammary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to yeast SUR4 membrane protein indicate that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of defects of fetal liver or defects of glucose-regulated ATPase activities.

FEATURES OF PROTEIN ENCODED BY GENE NO: 97

This gene is expressed primarily in fetal liver, brain, and amniotic fluid.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: defects of the fetal immune system and adult brain. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal immune system and adult brain, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein product of this clone is useful for detecting defects of the fetal immune and hematopoietic systems since fetal liver is

the predominant organ responsible for hematopoiesis in the fetus. In addition, the gene product of this gene is thought to be useful for detecting certain neurological defects of the brain.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 98

The translation product of this gene shares sequence homology with an yolk protein precursor, Vitellogenin which is thought to be important in binding lipids such as phosvitin.

This gene is expressed primarily in amniotic cells and fetal liver.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: defects in amniotic cells, fetal liver development and the fetal immune system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes
15 for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the [insert system where a related disease state is likely, e.g., immune], expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., amniotic cells, and liver, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma,
20 urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to vitellogenin indicate that the protein
25 product of this clone is useful for treatment and diagnosis of defects in amniotic cells, fetal liver development and the fetal immune system.

FEATURES OF PROTEIN ENCODED BY GENE NO: 99

This gene is expressed primarily in placenta, endometrial tumor, osteosarcoma
30 and stromal cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: tumor of the endometrium or bone, and osteosarcoma. Similarly, polypeptides and antibodies
35 directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the obstetric system (e.g. placenta,

endometrium) and the bones, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., placenta, endometrium, bone, and stromal cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of tumors and abnormalities of the endometrium, and the bones because of its abundance in the aforementioned tissues..

FEATURES OF PROTEIN ENCODED BY GENE NO: 100

This gene is expressed primarily in hepatocellular tumor.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: hepatocellular tumor. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the liver, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein product of this clone is useful for diagnosis and treatment of hepatocellular cancer because of its abundant expression in this tissue.

FEATURES OF PROTEIN ENCODED BY GENE NO: 101

This gene is expressed primarily in Corpus Colosum, fetal lung and infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: defects of the Corpus Colosum or defects of the fetal lung. Similarly, polypeptides and antibodies directed to

these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the Corpus Colosum and brain in general, and fetal lung, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., lung, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein product of this clone is useful for diagnosis and treatment of defects of the Corpus Colosum and brain in general, and defects of fetal lung.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 102

This gene is expressed primarily in T cells and stromal cells, and to a lesser extent in adrenal gland.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: defects of T cell immunity and stromal cell development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, stromal cells, and adrenal gland, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein product of this clone is useful for diagnosis and treatment of defects of T cell immunity and stromal cell development because of its abundant expression in these tissues.

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 103

This gene is expressed primarily in infant brain and placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: defects of the brain and nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, especially brain, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, and placenta, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein product of this clone is useful for detecting defects of the brain, especially in young children.

FEATURES OF PROTEIN ENCODED BY GENE NO: 105

This gene is expressed primarily in human osteoclastoma and to a lesser extent in human pancreas tumor.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer particularly osteoclastoma and pancreatic tumor. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly in transformed tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., bone and pancreas, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein product of this clone is useful for diagnosis and treatment of some types of tumors, particularly pancreatic cancer and osteoclastoma.

FEATURES OF PROTEIN ENCODED BY GENE NO: 106

This gene is expressed primarily in fetal liver/spleen, and to a lesser extent in activated T-Cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, spleen, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis or treatment of immune disorders.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 107

This gene is expressed primarily in human embryo and to a lesser extent in spleen and chronic lymphocytic leukemia.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: leukemia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune or hemopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., embryonic tissue, spleen, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein product of this clone is useful for the diagnosis and treatment of leukemia.

FEATURES OF PROTEIN ENCODED BY GENE NO: 108

This gene is expressed primarily in placenta, and to a lesser extent in early stage human brain and in lung.

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: fetal developmental abnormalities. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s)
10 or cell type(s). For a number of disorders of the above tissues or cells, particularly in fetal and amniotic tissue, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., placenta, brain and other tissue of the nervous system, and lung, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another
15 tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein product of this is useful for production of growth factor(s) associated with fetal development. Preferred
20 polypeptides comprise the full-length polypeptide shown in the sequence listing, truncated however, at the amino terminus and beginning with QTIE.

FEATURES OF PROTEIN ENCODED BY GENE NO: 109

This gene is expressed primarily in fetal spleen, and to a lesser extent in B-Cell
25 lymphoma and T-Cell lymphoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: lymphoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing
30 immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., spleen and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal
35 fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein product of this clone is useful for the treatment and diagnosis of human lymphomas.

FEATURES OF PROTEIN ENCODED BY GENE NO: 110

5 The translation product of this gene shares sequence homology with sarcoma amplified sequence (SAS), a tetraspan receptor which is thought to be important in malignant fibrous histiocytoma and liposarcoma.

This gene is expressed primarily in human osteoclastoma, and to a lesser extent in pineal gland and infant brain.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: malignant fibrous histiocytoma and liposarcoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification
15 of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., bone, pineal gland, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal
20 fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to sarcoma amplified sequence (SAS) indicate that the protein product of this clone is useful for treatment of, osteosarcoma,
25 malignant fibrous histiocytoma and liposarcoma and related cancers, particularly sarcomas.

FEATURES OF PROTEIN ENCODED BY GENE NO: 111

30 The translation product of this gene shares sequence homology with 6.8K proteolipid protein, mitochondrial - bovine.

This gene is expressed primarily in Wilm's tumor and to a lesser extent in cerebellum and placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
35 biological sample and for diagnosis of diseases and conditions: Wilm's tumor. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell

type(s). For a number of disorders of the above tissues or cells, particularly of the immune or renal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, and placenta, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to 6.8K proteolipid protein indicate that the protein product of this clone is useful for diagnostic and therapeutics associated with tumors, particularly Wilm's tumor disease.

FEATURES OF PROTEIN ENCODED BY GENE NO: 112

This gene is expressed primarily in embryonic tissue and to a lesser extent in osteoblasts, endothelial cells, macrophages (GM-CSF treated), and bone marrow.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., embryonic tissue, bone, endothelial cells, blood cells and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment or diagnosis of immune disorders. Preferred polypeptides encoded by this gene comprise the following amino acid sequence: MITDVQLAIFANMLGVSLFLLVVLVLYHYVAVNNPKKQE (SEQ ID NO: 636).

FEATURES OF PROTEIN ENCODED BY GENE NO: 113

This gene is expressed primarily in hepatocellular tumor, and to a lesser extent in fetal liver/spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: tumors, particularly hepatocellular tumors. Similarly, polypeptides and antibodies directed to these

5 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hepatic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, and spleen, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma,

10 urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein product of this clone is useful

15 for diagnosis and treatment of tumors, particularly hepatocellular tumors.

FEATURES OF PROTEIN ENCODED BY GENE NO: 114

The translation product of this gene exhibits a very high degree of sequence identity with the human Pig8 gene which is thought to be important in p53 mediated

20 apoptosis. The sequence of this gene has since been published by Polyak and colleagues (Nature 389, 300-306 (1997)). In addition, the predicted translation product of this contig exhibits very high sequence homology with a murine gene denoted as EI24 which is also thought to be important in p53 mediated apoptosis.

This gene is expressed primarily in infant brain and activated T-cells and to a

25 lesser extent in bone marrow, fetal liver, and prostate.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, and tissue damage by radiation and anti-cancer drugs. Similarly,

30 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the

35 nervous system, blood cells, bone marrow, liver, and prostate, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder,

relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to human Pig8 and murine EI24 genes indicate that polynucleotides and polypeptides corresponding to this gene are useful for preventing apoptosis in patients being treated with anti-oncogenic drugs such as etoposide, hydroperoxycyclophosphamide, and X-irradiation, since this protein product is upregulated in cells undergoing such treatment where p53 was overexpressed. It may also be useful in the treatment of hematopoietic disorders and in boosting numbers of hematopoietic stem cells by interfering with the apoptosis of progenitor cells. The mature polypeptide is predicted to comprise the following amino acid sequence:

EEMADSVKTFQLQDLARGIKDSIWGICTISKLDARIQQKREEQRRRRASSVLAQRRRAQSIERKQES
EPRIVSRIFQCCA WNGGVFWFSLLLFYRVFIPVLQSVTARIIGDPSLHGDVWSWLEFFLT SIFSA
LWVLPFLVLSKVVNAIWFQDIADLAFEVSGRKPFPFSVSKIADMLFNLLQALFLIQGMFVSL
FPIHLVGQLVSLHMSLLYSLYCFEYRWFNKGIEMHQRLSNIERNWPYYFGFGLPLAFLTAMQ
SSYIISGCLFSILFPLFIISANEAKTPGKAYLFQLRLFSLVVFLSNRLFHKTVYQLSALSSSTS AEK
FPSHPSPAKLKATAGH (SEQ ID NO: 637). Accordingly, polypeptides comprising the foregoing amino acid sequence are provided as are polynucleotides encoded such polypeptides.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 115

This gene is expressed primarily in stromal cells and to a lesser extent in multiple sclerosis.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: affecting the nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., stromal cells and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of multiple sclerosis and other autoimmune diseases.

FEATURES OF PROTEIN ENCODED BY GENE NO: 116

This gene is expressed primarily in the gall bladder

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: gall stones or infection of the digestive system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,
10 particularly of the digestive system or renal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., gall bladder and tissue of the digestive system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to
15 the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for possible prevention of digestive disorders where there may be a lack of digestive enzymes produced or in the detection and
20 possible prevention of gall stones.

FEATURES OF PROTEIN ENCODED BY GENE NO: 117

The translation product of this gene shares sequence homology with dystrophin gene which is thought to be important in building and maintenance of muscles.

25 This gene is expressed primarily in placenta and to a lesser extent in fetal brain and fetal liver, and spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: muscular dystrophy,
30 Duchenne and Becker's muscular dystrophies. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal muscle system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues
35 and cell types (e.g., placenta, brain and other tissue of the nervous system, muscle, liver, and spleen, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from

an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5 The tissue distribution and homology to the dystrophin gene indicate that polynucleotides and polypeptides corresponding to this gene are useful for diseases related the degenerative myopathies that are characterized by the weakness and atrophy of muscles without neural degradation; such as Duchenne and Becker's muscular dystrophies.

10 **FEATURES OF PROTEIN ENCODED BY GENE NO: 118**

This gene is expressed primarily in olfactory tissue and to a lesser extent in cartilage.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
15 biological sample and for diagnosis of diseases and conditions: connective tissue diseases; chondrosarcoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the connective tissue, expression of this gene at significantly higher or
20 lower levels may be routinely detected in certain tissues and cell types (e.g., olfactory tissue and cartilage, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the
25 disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for tumors of connective tissues, osteoarthritis and the treatment and diagnosis of chondrosarcoma.

30 **FEATURES OF PROTEIN ENCODED BY GENE NO: 119**

This gene is expressed primarily in Activated Neutrophils and to a lesser extent in fetal spleen, and CD34 positive cells from cord blood.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
35 biological sample and for diagnosis of diseases and conditions: allergies, defects in hematopoiesis and inflammation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential

identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and hematopoiesis system the, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, and spleen, and cancerous and
5 wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides
10 corresponding to this gene are useful for reducing the allergic effects felt by allergy suffers by neutralizing the activity of the immune system, especially since neutrophils are abundant in persons suffering from allergies and other inflammatory conditions.

FEATURES OF PROTEIN ENCODED BY GENE NO: 120

15 The translation product of this gene shares sequence homology with poly A binding protein II which is thought to be important in RNA binding for transcription of RNA to DNA

This gene is expressed primarily in colon and to a lesser extent in brain and immune system.

20 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: colon cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a
25 number of disorders of the above tissues or cells, particularly of the immune and digestive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., colon, tissue and cells of the immune system, and brain or other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal
30 fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to poly A binding protein II indicate that polynucleotides and polypeptides corresponding to this gene are useful for detection
35 and treatment of colon cancer and other disorders of the digestive system..

FEATURES OF PROTEIN ENCODED BY GENE NO: 121

The translation product of this gene shares sequence homology with thymidine diphosphoglucose 4.6 dehydrase which is thought to be important in the metabolism of sugar.

- 5 This gene is expressed primarily in fetal liver and spleen and to a lesser extent in infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: diabetes. Similarly,
10 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, spleen, and brain and other tissue of the
15 nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 20 The tissue distribution and homology to thymidine diphosphoglucose 4.6 dehydrase indicate that polynucleotides and polypeptides corresponding to this gene are useful for treatment of persons with diabetes since it appears that this protein is needed in the metabolism of sugar in to its more basic components.

25 **FEATURES OF PROTEIN ENCODED BY GENE NO: 122**

The translation product of this gene shares sequence homology with ceruloplasmin which is thought to be important in the metabolism and transport of iron and copper. Ceruloplasmin also contains domains with homology to clotting factors V and VIII. Defects in the circulating levels of ceruloplasmin (aceruloplasminemia) have
30 been associated with certain disease conditions such as Wilson disease, and the accompanying hepatolenticular degeneration.

This gene is expressed primarily in brain and retina and to a lesser extent in endothelial cells.

- Therefore, polynucleotides and polypeptides of the invention are useful as
35 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: diseases marked by defects in iron metabolism; aceruloplasminemia not characterized by defects in the

known ceruloplasmin gene locus; nonclassical Wilson disease; movement disorders; and tumors derived from a brain tissue origin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, retina, and nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, retinal tissue, and endothelial cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to ceruloplasmin indicate that polynucleotides and polypeptides corresponding to this gene are useful for treatment of patients with aceruloplasminemia, or other defects in iron and/or copper metabolism. Mutations in this locus could also be diagnostic for patients currently experiencing or predicted to experience aceruloplasminemia.

FEATURES OF PROTEIN ENCODED BY GENE NO: 123

This gene is expressed primarily in brain and B cell lymphoma and to a lesser extent in fetal liver and spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: B cell lymphoma; tumors and diseases of the brain and/or spleen; hematopoietic defects. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and hematopoietic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, blood cells, liver, and spleen, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment of disorders in neuronal,

hematopoietic, and immune systems. It could potentially be useful for neurodegenerative disorders and neuronal and/or hematopoietic cell survival or proliferation.

5 **FEATURES OF PROTEIN ENCODED BY GENE NO: 124**

This gene is expressed primarily in osteoclastoma, dermatofibrosarcoma, and B cell lymphoma and to a lesser extent in endothelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
10 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer in particular osteoclastoma, dermatofibrosarcoma, and B cell lymphoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of
15 the bone, immune, and circulatory system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., bone, epidermis, blood cells, and endothelial cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to
20 the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of cancers and lymphoma; osteoporosis; and the control of cell proliferation and/or differentiation.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 125

This gene is expressed primarily in immune tissues and hematopoietic cells, particularly in activated T cells and neutrophils, spleen, and fetal liver, and to a lesser extent in infant adrenal gland.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: defects in T cell activation; hematopoietic disorders; tumors of a hematopoietic and/or adrenal gland origin. Similarly, polypeptides and antibodies directed to these polypeptides are useful
30 in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic and/or endocrine systems, expression of this gene at significantly higher
35

- or lower levels may be routinely detected in certain tissues and cell types (e.g., cells and tissues of the immune system, hematopoietic cells, blood cells, liver, and adrenal gland, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual
- 5 having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for immune and/or hematopoietic disorders;
- 10 diseases related to proliferation and/or differentiation of hematopoietic cells; defects in T cell and neutrophil activation and responsiveness; and endocrine and/or metabolic disorders, particularly of early childhood.

FEATURES OF PROTEIN ENCODED BY GENE NO: 126

- 15 This gene is expressed primarily in placenta and endothelial cells and to a lesser extent in melanocytes and embryonic tissues.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: tumors of an endothelial
- 20 cell origin; angiogenesis associated with tumor development and metastasis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular system and developing embryo, expression of this gene at significantly higher or lower levels
- 25 may be routinely detected in certain tissues and cell types (e.g., placenta, endothelial cells, melanocytes, and embryonic tissues, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily
- 30 fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment of developmental disorders; inhibition of angiogenesis; and vascular patterning.

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 127

This gene is expressed primarily in endothelial cells and hematopoietic tissues, including spleen, tonsils, leukocytes, and both B- and T-cell lymphomas.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: tumors of an endothelial cell and/or hematopoietic origin; leukemias and lymphomas. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and vascular systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., endothelial cells, hematopoietic cells, spleen, tonsils, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the manipulation of angiogenesis; the differentiation and morphogenesis of endothelial cells; the proliferation and/or differentiation of hematopoietic cells; and the commitment of hematopoietic cells to distinct cell lineages.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 128

This gene is expressed primarily in kidney medulla and to a lesser extent in spleen from chronic myelogenous leukemia patients, prostate cancer, and some other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: tumors of a kidney origin; chronic myelogenous leukemia; prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the kidney and spleen, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., kidney, spleen, and prostate, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of kidney disorders and cancer, particularly chronic myelogenous leukemia and prostate cancer. It may also be useful for the enhancement of kidney tubule regeneration in the treatment of acute renal failure.

FEATURES OF PROTEIN ENCODED BY GENE NO: 129

This gene is expressed primarily in adult and infant brain and to a lesser extent in mesenchymal or fibroblast cells, as well as tissues with a mesenchymal origin.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: tumors of a brain and/or mesenchymal origin; neurodegenerative disorders; cancer; fibrosis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and of mesenchymal cells and tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis of tumors of a brain and/or mesenchymal origin; neurodegenerative disorders; cancer; and fibrosis, based upon the expression of this gene within those tissues. Fibrosis is considered as mesenchymal cells and fibroblasts are the primary cellular targets involved in this pathological condition.

FEATURES OF PROTEIN ENCODED BY GENE NO: 130

This gene is expressed primarily in hepatocellular cancer and to a lesser extent in fetal tissues as well as testes tumor.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: liver cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing

immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, fetal tissue, and testes and other
5 reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of liver cancer.

FEATURES OF PROTEIN ENCODED BY GENE NO: 131

This gene is expressed only in infant early brain.

15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: development and diseases of the nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification
20 of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another
25 tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating diseases of the brain in children and in
30 treating nervous system disorders such as Alzheimer's disease, schizophrenia, dementia, depression, etc.

FEATURES OF PROTEIN ENCODED BY GENE NO: 132

This gene is expressed primarily in brain and to a lesser extent in glioblastoma.

35 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: Alzheimer's disease,

schizophrenia, depression, mania, and dementia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating brain disorders such as Alzheimer's disease, schizophrenia, depression, mania, and dementia.

15 **FEATURES OF PROTEIN ENCODED BY GENE NO: 133**

The translation product of this gene shares sequence homology with ribitol dehydrogenase of bacteria which is thought to be important in metabolism of sugars.

This gene is expressed primarily in macrophage and to a lesser extent in T-cell lymphoma and lung.

20 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: tissue destruction in inflammation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells and lung, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30 The tissue distribution and homology to ribitol dehydrogenase indicate that polynucleotides and polypeptides corresponding to this gene are useful for altering macrophage metabolism in diseases such as inflammation where macrophages are causing excess tissue destruction.

FEATURES OF PROTEIN ENCODED BY GENE NO: 134

This gene is expressed primarily in pancreatic tumor and to a lesser extent in synovial sarcoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to,. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine and connective tissue systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., pancreas, and synovial tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating and diagnosing various cancers.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 135

This gene is expressed primarily in T cell lines such as Raji and to a lesser extent in infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: immune system disorders and inflammation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating and diagnosing inflammatory diseases

such as rheumatoid arthritis, sepsis, inflammatory bowel disease, and psoriasis, as well as neutropenia.

FEATURES OF PROTEIN ENCODED BY GENE NO: 136

5 The translation product of this gene shares high sequence homology with SAR1 subfamily of GTP-binding proteins which is thought to be important in vesicular transport in mammalian cells.

 This gene is expressed primarily in serum-stimulated smooth muscle cells and to a lesser extent in a T-cell lymphoma.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: diseases affecting vesicular transport. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification
15 of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the muscular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, and smooth muscle, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample
20 taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 The tissue distribution and homology to GTP-binding proteins indicate that polynucleotides and polypeptides corresponding to this gene are useful for gene therapy
25 in treating the large number of diseases involved in defective vesicular transport within cells..

FEATURES OF PROTEIN ENCODED BY GENE NO: 137

 The translation product of this gene shares sequence homology with a protein
30 found in *C. elegans* cosmid F25B5.

 This gene is expressed primarily in a fetal tissues and to a lesser extent in melanocytes.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
35 biological sample and for diagnosis of diseases and conditions: abnormal fetal development, especially of the pulmonary system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes

for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal pulmonary system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., fetal tissue, pulmonary tissue, and melanocytes, and
5 cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides
10 corresponding to this gene are useful for treatment and diagnosis of diseases affecting the pulmonary system, such as emphysema.

FEATURES OF PROTEIN ENCODED BY GENE NO: 138

This gene is expressed primarily in gall bladder and to a lesser extent in smooth
15 muscle.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: digestive system disease and gall bladder problems. Similarly, polypeptides and antibodies directed to these
20 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., gall bladder and tissue of the digestive system, and smooth muscle, and cancerous and
25 wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides
30 corresponding to this gene are useful for treating diseases of the digestive system.

FEATURES OF PROTEIN ENCODED BY GENE NO: 139

This gene is expressed primarily in placenta and to a lesser extent in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as
35 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: abnormal fetal development. Similarly, polypeptides and antibodies directed to these polypeptides are

useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of developing tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., placenta, and brain and other
5 tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating and diagnosing abnormal fetal development.

FEATURES OF PROTEIN ENCODED BY GENE NO: 140

15 This gene is expressed primarily in smooth muscle and to a lesser extent in ovary, prostate cancer, and activated monocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: hypertension and
20 atherosclerosis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the circulatory system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., smooth
25 muscle, ovary and other reproductive tissue, prostate, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating diseases of the circulatory system, such as hypertension, atherosclerosis, etc.

FEATURES OF PROTEIN ENCODED BY GENE NO: 141

35 This gene is expressed primarily in fetal spleen and to a lesser extent in placenta and bone marrow.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: anemia and other diseases affecting blood cells. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the circulatory and pulmonary systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., spleen, placenta, bone marrow, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the generation of red and white blood cells and for the diagnosis of disease of these cells.

FEATURES OF PROTEIN ENCODED BY GENE NO: 142

The predicted translation product of this contig is a human homolog of the murine tetracycline/sugar transporter molecule recently reported by Matsuo and colleagues (Biochem. Biophys. Res. Commun. 238 (1), 126-129 (1997)).

This gene is expressed primarily in synovium and to a lesser extent in endothelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: rheumatoid arthritis and inflammation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and lymphatic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., synovial tissue, and endothelial cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of inflammatory diseases, such as rheumatoid arthritis, leukemia, neutropenia, inflammatory bowel disease, psoriasis, sepsis, and the like.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 143

This gene is expressed primarily in placenta and to a lesser extent in melanocyte, fetal liver and spleen, and bone marrow.

Therefore, polynucleotides and polypeptides of the invention are useful as
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: abnormal early development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, lower levels
15 may be routinely detected in certain tissues and cell types (e.g., placenta, melanocytes, liver, spleen, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an
20 individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of abnormal early development phenomena and diseases.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 144

This gene is expressed primarily in fetal liver and spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: anemia and neutropenia.
30 Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and blood systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver and spleen,
35 and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the

expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful in hematopoiesis and bone marrow regeneration as it is most abundant in fetal tissues responsible for the generation of hematopoietic cells.

FEATURES OF PROTEIN ENCODED BY GENE NO: 145

The translation product of this gene shares sequence homology with protein tyrosine phosphatase which is thought to be important in transducing signal to activate cells such as T cell, B cell and other cell types.

This gene is expressed primarily in T cells and tissues in early stages of development and to a lesser extent in cancers.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: immuno-related diseases and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., embryonic and fetal tissue, undifferentiated cells, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to the protein tyrosine phosphatase family indicate that polynucleotides and polypeptides corresponding to this gene are useful for modulating the immune system.

FEATURES OF PROTEIN ENCODED BY GENE NO: 146

This gene is expressed primarily in T cell and to a lesser extent in B cell, macrophages and tumor tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: immuno-disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in

providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for regulating the immune system therefore can be used in treating diseases such as autoimmune diseases and cancers.

FEATURES OF PROTEIN ENCODED BY GENE NO: 147

This gene is expressed primarily in placenta and to a lesser extent in endothelial cells, testis tumor, ovarian cancer, uterine cancer.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., placenta, endothelial cells, testis and ovary and other reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of cancers.

FEATURES OF PROTEIN ENCODED BY GENE NO: 148

This sequence has significant homology to mouse torsin A. Recently, another group cloned the human Torsin A gene. (See, Accession No. 2358279; see also Nature Genet. 17, 40-48 (1997).)

This gene is expressed primarily in osteoclastoma, T-cell, and placenta and to a lesser extent in fetal lung, fetal liver, fetal brain, adult brain and tumor tissues

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: disease conditions in hematopoiesis and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoiesis system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, bone, placenta, lung, liver, and brain and other tissues of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treating blood related diseases such as deficiencies in red blood cell, white blood cell, platelet and other hematopoiesis cells.

FEATURES OF PROTEIN ENCODED BY GENE NO: 149

This gene is expressed primarily in T cell, prostate and prostate cancer, endothelial cells and to a lesser extent in monocyte, dendritic cell, bone marrow, salivary gland, colon cancer, stomach cancer, pancreatic tumor, uterine cancer, fetal spleen and osteoclastoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: immuno-related diseases and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, prostate, endothelial cells, dendritic cells, bone marrow, salivary gland, colon, stomach, pancreas, uterus, spleen and bone, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment of cancers.

FEATURES OF PROTEIN ENCODED BY GENE NO: 150

- 5 This gene was recently cloned by another group, calling it eIF3-p66. (See Accession No. 2351378.) This gene plays a role in RNA binding and macromolecular assembly, and therefore, any mutations in this gene would likely result in a diseased phenotype. Preferred polypeptide fragments comprise the amino acid sequence:
- 10 MAKFMTPVIQDNPSGWGPCAVPEQFRDMPYQPFSGDRLGKVADWTGATYQDKRYTNKYSS
QFGGGSQYAYFHEEDESSFLVDTARTQKTAYQRNRMRFQARNLRRDKDRRNMLQFNLQILP
KSAKQKERERIRLQKKFQKQFGVRQKWDQKSQKPRDSSVEVRSDWEVKEEMDFPQLMKMRY
LEVSEPQDIECCGALEYDYKAFDRITRSEKPLRXXKRIFHTVTTTDDPVIRKLAKTQGNVFATD
AILATLMSCTRSVYSWDIVVQRVGSKLFFDKRDNSDFDLLTVSETANEPPQDEGNSFNSPRNL
AMEATYINHNFSSQCLRMGKERYNFPNPNPFVEDDMDKNEIASVAYRYSRSGKLGDIDLIVRC
15 EHDGVMTGANGEVSFINIKTLNEWDSRHCNGVDWRQKLD SQRGAVIATELKNN SYKLARWTC
CALLAGSEYLKLGYSRYHVKDSSRHVILGTQQFKPNEFASQINLSVENAWGILRCVIDICMKL
EEGKYLLKDPNKQVIRVYSLPDGTFSS (SEQ ID NO: 638), as well as N-terminal and C-terminal deletions of this polypeptide fragment.

- 20 This gene is expressed primarily in T cell, bone marrow, embryo and endothelial cells and to a lesser extent in testis tumor and endometrial tumor.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: immune diseases and tumors. Similarly, polypeptides and antibodies directed to these polypeptides are useful
- 25 in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial
- 30 fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for immune disorders and cancers.

FEATURES OF PROTEIN ENCODED BY GENE NO: 151

This gene is expressed primarily in testis and to a lesser extent in T cell, spinal cord, placenta, neutrophil and monocyte.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: male reproductive and endocrine disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive, immune and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., testis and other reproductive tissue, blood cells, tissue of the nervous system, and placenta, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for regulating immune and reproductive functions.

FEATURES OF PROTEIN ENCODED BY GENE NO: 152

The translation product of this gene shares sequence homology with tyrosyl-tRNA synthetase which is thought to be important in cell growth.

This gene is expressed primarily in brain, liver, keratinocytes, tonsils, and heart.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer autoimmune diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, liver, keratinocytes, tonsils, heart expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissues of the nervous system, liver, keratinocytes, tonsils and heart, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard

gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to tyrosyl-tRNA synthetase indicate that polynucleotides and polypeptides corresponding to this gene are useful for modulating cell growth.

FEATURES OF PROTEIN ENCODED BY GENE NO: 153

This gene is homologous to the *Drosophila* transcriptional regulator dre4. (See Accession No. 2511745.) Dre4 is a gene required for steroidogenesis in *Drosophila melanogaster* and encodes a developmentally expressed homologue of the yeast transcriptional regulator CDC68. Preferred polypeptide fragments comprise the amino acid sequence: KKRHTDVQFYTEVGEITDLGKHQMHDRDDLYAEQMEREMRHKLKTAFAKNFIEKVEALTKEELEFEVPRDLGFNGAPYRSTCLLQPTSSALVNATEWPPFVVTLDEVELIHFXRVQFHLKNFDMVIVYKDYSKKVTMINAIPVASLDPIKEWLNSCDLKYTEGVQSLNWTIKMTIVDDPEGFFEQGGWSFL (SEQ ID NO: 639), as well as N-terminal and C-terminal deletions of this fragments. Also preferred are polynucleotide fragments encoding this polypeptide fragment.

This gene is expressed primarily in fetal liver, spleen, placenta, lung, T cell, thyroid, testes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: brain tumor, heart and liver diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal liver, spleen, placenta, lung, T cell, thyroid, testes expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, spleen, placenta, lung, blood cells, thyroid, and testes and other reproductive tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

FEATURES OF PROTEIN ENCODED BY GENE NO: 154

This gene is expressed primarily in brain and to a lesser extent in fetal heart, testis, spleen, lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: heart, liver and spleen diseases, immunological diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, fetal heart, testis, spleen, lung expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, heart, testes and other reproductive tissue, spleen, and lung, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 155

Activation of T cells through the T cell antigen receptor (TCR) results in the rapid tyrosine phosphorylation of a number of cellular proteins, one of the earliest being a 100 kDa protein. This gene is the human equivalent of murine valosin containing protein (VCP). VCP is a member of a family of ATP binding, homo-oligomeric proteins, and the mammalian homolog of *Saccharomyces cerevisiae* cdc48p, a protein essential to the completion of mitosis in yeast. Both endogenous and expressed murine VCP are tyrosine phosphorylated in response to T cell activation. Thus we have identified a novel component of the TCR mediated tyrosine kinase activation pathway that may provide a link between TCR activation and cell cycle control.

20

This gene is expressed primarily in brain, liver, spleen, placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer immunological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, liver, spleen, placenta expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, liver, spleen, and placenta, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from

30

35

an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5 The tissue distribution and homology to VCR indicate that polynucleotides and polypeptides corresponding to this gene are useful for treating cancer.

FEATURES OF PROTEIN ENCODED BY GENE NO: 156

10 The translation product of this gene shares sequence homology with rat growth response protein which is thought to be important in cell growth. A group recently cloned the human homolog of this gene, calling it insulin induced protein 1. (See Accession No. 2358269, see also, Genomics 43 (3), 278-284 (1997).) Preferred polypeptide fragments comprise the amino acid sequence: RSGLGLGITIAFLATLITQF LVYNGVYQYTSPDFLYIRSWLPCIFFSGGVTVGNIGRQLAMGVPEKPHSD (SEQ ID NO: 640), as well as N-terminal and C-terminal deletions of this polypeptide fragment. Also
15 preferred are polynucleotide fragments encoding these polypeptide fragments.

This gene is expressed primarily in brain, liver, placenta, heart, spleen, lymphoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
20 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer immunological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, liver, placenta, heart, spleen.
25 expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, liver, placenta, heart, spleen, and lymphoid tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the
30 standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to growth-response protein indicate that polynucleotides and polypeptides corresponding to this gene are useful for modulating cell growth.

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 157

This gene is expressed primarily in Glioblastoma, endometrial tumor, lymphoma and pancreas tumor.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: Glioblastoma, Endometrial tumor, lymphoma and pancreas tumor. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., endometrium, lymphoid tissue, pancreas, and tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

FEATURES OF PROTEIN ENCODED BY GENE NO: 158

The translation product of this gene shares sequence homology with IGE receptor which is thought to be important in allergy and asthma.

This gene is expressed primarily in T cell, and fetal liver.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: allergy and asthma and other immunological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, and liver, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to IgE receptor indicate that polynucleotides and polypeptides corresponding to this gene are useful for allergy and asthma.

5 **FEATURES OF PROTEIN ENCODED BY GENE NO: 159**

The translation product of this gene shares sequence homology with immunoglobulin heavy chain which is thought to be important in immune response to the antigen.

10 This gene is expressed primarily in activated neutrophil and to a lesser extent in activated T cell, monocyte and heart.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: infection, inflammation and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are
15 useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, and heart, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial
20 fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to immunoglobulin heavy chain variable region indicate that polynucleotides and polypeptides corresponding to this gene are
25 useful for making the ligand to block specific antigen which cause certain disease.

FEATURES OF PROTEIN ENCODED BY GENE NO: 160

The translation product of this gene shares sequence homology with mouse X inactive specific transcript protein which is thought to be important in X chromosome
30 inactivation.

This gene is expressed primarily in HSA172 cell and to a lesser extent in normal ovary tissue, ovarian cancer, frontal cortex and brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
35 biological sample and for diagnosis of diseases and conditions: ovarian tumor, schizophrenia and other neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for

differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and neural system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., ovary and other reproductive tissue, and brain and other
5 tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10 The tissue distribution and homology to X inactive specific transcript protein indicate that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of reproductive system tumors and CNS tumors.

FEATURES OF PROTEIN ENCODED BY GENE NO: 161

15 This gene is expressed primarily in adipose cell and to a lesser extent in liver and prostate.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: obesity and liver
20 disorder. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the adipose cell, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., adipose cells, liver, and
25 prostate, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30 The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment of obesity and liver disorder.

FEATURES OF PROTEIN ENCODED BY GENE NO: 162

35 The translation product of this gene shares sequence homology with yeast ubiquitin activating enzyme homolog which is thought to be important in protein posttraslation processing.

This gene is expressed primarily in stromal cell and to a lesser extent in retina, H. Atrophic Endometrium, colon carcinoma and myeloid progenitor cell.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: defects of stromal cell development, neuronal growth disorders and tumors. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and neural system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., retinal cells, endometrium, colon, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to ubiquitin-activating enzyme homolog indicate that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis or treatment of some type of tumors, fucosidosis and neuronal growth disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 163

This gene is expressed primarily in primary breast cancer and hemangiopericytoma and to a lesser extent in adult brain and cerebellum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: breast cancer, leukemia and cerebellum disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and neural system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., mammary tissue, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis or treatment of various tumors and disease involved in neural system.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 164

The translation product of this gene shares sequence homology with proline rich proteins. Recently, another group has also cloned this gene, calling it CD84 leukocyte antigen, a new member of the Ig superfamily. (See Accession No. U82988, see also, Blood 90 (6), 2398-2405 (1997).)

- 10 This gene is expressed primarily in Weizmann olfactory tissue and osteoclastoma and to a lesser extent in anergic T-cell.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: osteoporosis and immune
15 disease. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., olfactory tissue, bone, and
20 blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 25 The tissue distribution and homology to the Ig superfamily indicate that the protein product of this clone is useful for treatment of osteoporosis, autoimmune disease, and other immune disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 165

- 30 This gene is expressed primarily in atrophic endometrium and colon cancer and to a lesser extent in some fetal tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: tumors. Similarly,
35 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system,

expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., endometrium, colon, and fetal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of tumors, specifically endometrium and colon tumors.

FEATURES OF PROTEIN ENCODED BY GENE NO: 166

This gene is expressed primarily in human primary breast cancer and to a lesser extent in activated monocyte. Although the predicted signal sequence is identified in Table 1, other upstream sequences are also relevant. Preferred polypeptide fragments comprise the amino acid sequence: VTQPKHLSASMGGSEIPFSFYYPWELAXXPXVRISWRRGHFHG QSFYSTRPPSIHKDYVNRLFLNWTEGQESGFLRISNLRKEDQSVYFCRVELDTRRSG (SEQ ID NO: 641), as well as N-terminal and C-terminal deletions. Also preferred are polynucleotide fragments encoding these polypeptide fragments.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: breast cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., mammary tissue, and blood cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis of breast cancer.

FEATURES OF PROTEIN ENCODED BY GENE NO: 167

This gene is expressed primarily in fetal tissues and to a lesser extent in adult lung. This gene has also been mapped to chromosomal location 9q34, and thus, can be used as a marker for linkage analysis for chromosome 9.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the embryo tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., fetal tissues, and lung, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

FEATURES OF PROTEIN ENCODED BY GENE NO: 168

The translation product of this gene shares sequence homology with Ig Heavy Chain which is thought to be important in immune response.

This gene is expressed primarily in prostate cancer tissue specifically

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the prostate, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., prostate, tissue and cells of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

FEATURES OF PROTEIN ENCODED BY GENE NO: 169

The translation product of this gene shares sequence homology with cytosolic acyl coenzyme-A hydrolase, which is thought to be important in neuron-specific fatty acid metabolism. The gene represented by this contig has since been published by Hajra and colleagues (GenBank Accession No. U91316).

This gene is expressed primarily in human pituitary gland and to a lesser extent in colorectal cancer tissue. This gene has also been observed in the LNCAP cell line.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: hyperlipidemias of familial and/or idiopathic origins. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly blood, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., pituitary and colon, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to rat cytosolic acyl coenzyme-A hydrolase indicate that polynucleotides and polypeptides corresponding to this gene are useful for the detection or treatment of hyperlipidemia disease states by virtue of the ability of specific drugs to activate the enzyme.

FEATURES OF PROTEIN ENCODED BY GENE NO: 170

The translation product of this gene shares sequence homology with a *Caenorhabditis elegans* gene which is thought to be important in organism development.

This gene is expressed primarily in human synovial sarcoma tissue, bone marrow, and to a lesser extent in human brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, of bone, specifically synovial sarcoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the bone, connective tissues and possibly immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., synovial tissue, bone marrow, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another

tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 The tissue distribution and homology to *Caenorhabditis elegans* indicate that polynucleotides and polypeptides corresponding to this gene are useful as a diagnostic and/or therapeutic modality directed at the detection and/or treatment of connective tissue sarcomas or other related bone diseases.

FEATURES OF PROTEIN ENCODED BY GENE NO: 171

- 10 The translation product of this gene shares sequence homology with beta1-6GlcNAc transferase which is thought to be important in the transfer and metabolism of beta1-6, N-acetylglucosamine. This gene product has previously been shown to suppress melanoma lung metastasis in both syngeneic and nude mice, decreased invasiveness into the matrigel, and inhibition of cell attachment to collagen and laminin
15 without affecting cell growth.

This gene is expressed primarily in human testes and prostate tissues, and to a lesser extent in kidney, medulla, and pancreas.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
20 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer particularly melanoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at
25 significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., testes and other reproductive tissue, prostate, kidney, pancreas, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard
30 gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- The tissue distribution and homology to beta1-6GlcNAc transferase indicate that the protein product of this clone is useful for the development of diagnostic and/or therapeutic modalities directed at the detection and/or treatment of cancer, the metastasis
35 of malignant tissue or cells. Defects in this potentially secreted enzyme may play a role in metastasis.

FEATURES OF PROTEIN ENCODED BY GENE NO: 172

This gene is expressed primarily in fetal spleen and liver.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions: immune disorders,
Wilm's tumor disease, hepatic disorders, and hematopoietic disorders. Similarly,
polypeptides and antibodies directed to these polypeptides are useful in providing
immunological probes for differential identification of the tissue(s) or cell type(s). For a
10 number of disorders of the above tissues or cells, particularly of the hematopoiesis and
immune systems, expression of this gene at significantly higher or lower levels may be
routinely detected in certain tissues and cell types (e.g., spleen and liver, and cancerous
and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or
spinal fluid) or another tissue or cell sample taken from an individual having such a
15 disorder, relative to the standard gene expression level, i.e., the expression level in
healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides
corresponding to this gene are useful for the treatment and identification of fetal defects
along with correcting diseases that affect hematopoiesis and the immune system.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 173

The translation product of this gene shares sequence homology with ret II
oncogene which is thought to be important in Hirschsprung disease and many types of
cancers.

25 This gene is expressed in multiple tissues including the lymphatic system, brain,
and thyroid.

Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for identification of the tissue(s) or cell type(s) present in a biological sample
and for diagnosis of diseases and conditions: Hirschsprung disease and multiple
30 cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful
in providing immunological probes for identification of the tissue(s) or cell type(s). For
a number of disorders of the above tissues or cells, particularly of the immune and
central nervous system, expression of this gene at significantly higher or lower levels
may be routinely detected in certain tissues and cell types (e.g., lymphoid tissue,
35 thyroid, and brain and other tissue of the nervous system, and cancerous and wounded
tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or
another tissue or cell sample taken from an individual having such a disorder, relative to

the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to ret II oncogene indicate that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of various cancers. It would also be useful for the diagnosis and treatment of Hirschsprung disease. Preferred polypeptides of the invention comprise the amino acid sequence: MEAQQVNEAESAREQLQXLHDQIAGQKASKQELETelerLKQEFHYIEEDLY RTKNTLQSRDKDRDEEIQKLRNQLTNKTLSSSSQSELENRLHQLTETLIQKQTMLESLSSTEKNSL VFQLERLEQQMNSASGSSSSNGSSINMSGIDNGEGTRLRNVPVLFNDTETNLAGMYGKVRKAAS
10 SIDQFSIRLGIFLRRYPIARVFVHIYMALLHLWVMIVLLTYTPEM HHDQPYGK (SEQ ID NO: 642).

FEATURES OF PROTEIN ENCODED BY GENE NO: 174

The translation product of this gene shares sequence homology with testis enhanced gene transcript which is thought to be important in regulation of human development.

This gene is expressed primarily in infant brain and to a lesser extent in a variety of other tissues and cell types, including the prostate, testes, monocytes, macrophages, dendritic cells, keratinocytes, and adipocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: neurological, developmental, immune and inflammation disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes
25 for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, prostate, testes and other reproductive tissue, blood cells, keratinocytes, and adipocytes, and
30 cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to testis enhanced gene transcript indicate
35 that the protein product of this clone is useful for diagnosis and treatment of disorders involving the developing brain and the immune system.

FEATURES OF PROTEIN ENCODED BY GENE NO: 175

This gene is expressed primarily in prostate and to a lesser extent in various other tissues, including placenta.

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancers, especially of the prostate. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for
10 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the prostate, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., prostate and placenta, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell
15 sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein product of this clone is useful for diagnosis and treatment of prostate disorders and cancer. It may also be useful for
20 the diagnosis and treatment of endocrine disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 176

The translation product of this gene shares sequence homology with *Sacchromyces cerevisiae* YNT20 gene which is thought to be important in
25 mitochondrial function.

This gene is expressed at a particularly high level in muscle tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases related to such tissues and cell types
30 including: muscle wasting diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neuromuscular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell
35 types (e.g., muscle and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,

the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to the YNT20 gene indicate that this protein is useful for treatment and detection of neuromuscular diseases caused by loss of mitochondrial function. For example this gene or its protein product could be used in replacement therapy for such diseases.

FEATURES OF PROTEIN ENCODED BY GENE NO: 177

This gene is expressed primarily in the brain and to a lesser extent in kidney, placenta, smooth muscle, heart and lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: neuromuscular diseases, degenerative diseases of the central nervous system, and heart disease. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neuromuscular system, central nervous system, and heart, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, kidney, placenta, muscle, heart and lung, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

This gene or its protein product could also be used for replacement therapy for the above mentioned diseases.

FEATURES OF PROTEIN ENCODED BY GENE NO: 178

The translation product of this gene shares sequence homology with caldesmon which is thought to be important in the cellular response to changes in glucose levels.

This gene is expressed primarily in multiple tissues including brain and retina.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: central nervous system disorders and retinopathy. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for identification of the tissue(s) or cell

type(s). For a number of disorders of the above tissues or cells, particularly of the CNS disorders and retinopathy, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, and retinal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to caldesmon indicate that polynucleotides and polypeptides corresponding to this gene are useful for treatment of retinopathies.

FEATURES OF PROTEIN ENCODED BY GENE NO: 179

The translation product of this gene shares sequence homology with mouse fibrosin protein which is thought to be important in regulation of fibrinogenesis in certain chronic inflammatory diseases.

This gene is expressed primarily in amniotic cells and breast tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of breast cancer and abnormal embryo development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., amniotic cells, and mammary tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to fibrosin indicate that the protein product of this clone is useful for treatment of breast cancer. This gene or its protein product could be used in replacement therapy for breast cancer. In addition the protein product of this gene is useful in the treatment of chronic inflammatory diseases.

FEATURES OF PROTEIN ENCODED BY GENE NO: 180

This gene is expressed several infant tissues including brain and liver and various adult tissues including brain, lung, liver, testes, and prostate.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, brain cancer, lung cancer, liver cancer and cancers of the reproductive system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, hepatic system, and reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, lung, liver, testes and other reproductive tissue, and prostate, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of this gene product indicates that the protein product of this clone is involved in growth regulation and could be used as a growth factor or growth blocker in a variety of settings including treatment of cancers.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 181

This gene is expressed primarily in activated monocytes and to a lesser extent in melanocytes and dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of immune system diseases and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., blood cells, melanocytes, and dendritic cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein product of this clone could be involved in growth regulation and could be used as a growth factor or growth blocker in a variety of settings including treatment of cancers.

FEATURES OF PROTEIN ENCODED BY GENE NO: 182

This gene is expressed primarily in placenta and several tumors of various tissue origin and to a lesser extent in normal tissues including liver, lung, brain, and skin,

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of cancers of all kinds. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders
10 of the above tissues or cells, particularly of the central nervous system, respiratory system and skin, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., liver, lung, brain and other tissues of the nervous system, and skin, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or
15 cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The high expression of this gene in multiple tumors indicates that the protein product of the clone may be involved in cell growth control and therefore would be
20 useful for treatment of certain cancers. Likewise molecules developed to block the activity of the protein product of this clone could be used to block its potential role in tumor growth promotion.

FEATURES OF PROTEIN ENCODED BY GENE NO: 183

25 The translation product of this gene shares sequence homology with the mouse Ndr1 gene which is thought to be important in cancer progression.

This gene is expressed multiple cell types and tissues including brain, lung, kidney, bone marrow, liver, and spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as
30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of all types of cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous, immune, and endocrine
35 systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, lung, kidney, bone marrow, liver and spleen, and cancerous and wounded

tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 The tissue distribution and homology to Ndr1 gene, which is thought to be involved in cancer progression, indicate that polynucleotides and polypeptides corresponding to this gene are useful for treatment of certain cancers. Likewise molecules developed to block the activity of the protein product of this clone could be used to block its potential role in tumor growth promotion.

10

FEATURES OF PROTEIN ENCODED BY GENE NO: 184

This gene is expressed primarily in early stage human brain and liver and to a lesser extent in several other fetal tissues.

- 15 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions: brain and liver cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the
- 20 central nervous system and immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., brain and other tissue of the nervous system, liver, and fetal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder,
- 25 relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The expression of this gene in embryonic tissues indicates that the protein could be involved in growth regulation and could be used as a growth factor or growth blocker in a variety of settings including treatment of cancers.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 185

This gene is expressed primarily in infant and embryonic brain.

- 35 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of degenerative nervous system disorders and brain cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell

type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., embryonic tissue, brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The expression of this gene in embryonic tissues indicates that the protein could be involved in growth regulation and could be used as a growth factor or growth blocker in a variety of settings including treatment of cancers.

FEATURES OF PROTEIN ENCODED BY GENE NO: 186

This gene is expressed primarily in multiple tissues including placenta, fetal lung, fetal liver, and brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of all types of cancers including liver, brain and lung. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, pulmonary system, and hepatic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues and cell types (e.g., placenta, lung, liver, and brain and other tissue of the nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The expression of this gene in embryonic tissues indicates that the protein could be involved in growth regulation and could be used as a growth factor or growth blocker in a variety of settings including treatment of cancers.

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
1	HTTEZ21	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	11	582	1	582	177	177	313	1	18	19	22
1	HTTEZ21	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	197	1020	296	830	442	442	499	1	18	19	22
2	HBGBW52	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	12	465	1	465	81	81	314	1	30	31	128
2	HBGBW52	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	198	524	229	343		196	500	1	20	21	33
3	HCUFM41	97897 02/26/97 209043 05/15/97	ZAP Express	13	474	1	474	1	1	315	1	24	25	28
3	HCUFM41	97897 02/26/97 209043 05/15/97	ZAP Express	199	332	1	319	35	35	501	1	24	25	28

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
4	HCUFQ22	97897 02/26/97 209043 05/15/97	ZAP Express	14	314	1	298	122	122	316	1	34	35	64
5	HCUFV01	97897 02/26/97 209043 05/15/97	ZAP Express	15	613	1	613	30	30	317	1	18	19	21
6	HCUGA50	97897 02/26/97 209043 05/15/97	ZAP Express	16	356	1	356	239	239	318	1	22	23	39
7	HCUIM14	97897 02/26/97 209043 05/15/97	ZAP Express	17	414	185	414	278	278	319	1	26	27	33
8	HLDOU93	97897 02/26/97 209043 05/15/97	pCMVSPORT 3.0	18	469	1	469	77	77	320	1	44	45	88
9	HEIAX07	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	19	550	1	550	129	129	321	1	21	22	23

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
9	HEIAX07	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	200	376	9	376		1	502	1	8	9	15
10	HSAXR76	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	20	741	55	741	190	190	322	1			27
11	HNGJJ68	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	21	991	1	991	62	62	323	1	30	31	64
11	HNGJJ68	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	201	1192	253	1137		409	503	1			19
12	HCFAW04	97897 02/26/97 209043 05/15/97	pSport1	22	653	1	653	64	64	324	1	30	31	196
12	HCFAW04	97897 02/26/97 209043 05/15/97	pSport1	202	589	1	513	109	109	504	1			29

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
13	HLMAV65	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	23	1486	596	1418	102	102	325	1	54	55	252
13	HLMAV65	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	203	847	1	839	87	87	505	1	30	31	75
13	HLMAV65	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	204	852	75	850		690	506	1			10
13	HTXEF04	209235 09/04/97	Uni-ZAP XR	205	1354	54	1354	100	100	507	1	33	34	207
14	HPMFD84	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	24	2323	1017	2059	1242	1242	326	1	21	22	68
14	HPMFD84	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	206	1378	113	1226	303	303	508	1	25	26	36
15	HE6DB26	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	25	683	1	683	304	304	327	1	30	31	84

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
15	HE6DB26	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	207	1166	281	884	567	567	509	1	18	19	19
16	HHFFL33	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	26	2036	14	1959	214	214	328	1	20	21	36
17	HODBD33	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	27	717	1	717	70	70	329	1	30	31	63
17	HODBD33	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	208	697	2	697	33	33	510	1	31	32	32
18	HMDAE90	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	28	495	1	495	39	39	330	1	24	25	35
19	HOUAW01	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	29	556	1	556	116	116	331	1	19	20	23

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
20	HBIAE44	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	30	434	1	434	78	78	332	1	35	36	40
21	HCFME41	97897 02/26/97 209043 05/15/97	pSport1	31	715	1	715	87	87	333	1	30	31	111
21	HCFME41	97897 02/26/97 209043 05/15/97	pSport1	209	932	274	932	387	387	511	1	27	28	28
22	HOGCO71	97897 02/26/97 209043 05/15/97	pCMVSPORT 2.0	32	486	1	486	137	137	334	1	21	22	106
23	HOSEX08	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	33	725	1	725	436	436	335	1	30	31	50
23	HOSEX08	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	210	661	1	647	81	81	512	1	25	26	26

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
24	HSKNJ72	97897 02/26/97 209043 05/15/97	pBluescript	34	437	1	437	85	85	336	1	30	31	48
25	HEBEB69	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	35	943	1	943	196	196	337	1	30	31	41
25	HEBEB69	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	211	592	1	534	72	72	513	1	24	25	33
26	HE6EH18	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	36	604	1	604	375	375	338	1	20	21	76
26	HE6EH18	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	212	938	1	509		17	514	1	30	31	47
27	HSAUZ47	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	37	349	1	349		201	339	1	20	21	31

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
28	HSSDM73	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	38	672	1	672	22	22	340	1	38	39	42
29	HBMVK68	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	39	1908	135	1908	309	309	341	1	20	21	26
30	HMKDC66	97898 02/26/97 209044 05/15/97	pSport1	40	458	93	458	147	147	342	1	24	25	26
31	HMKCU94	97898 02/26/97 209044 05/15/97	pSport1	41	1153	500	1153	427	427	343	1	30	31	157
31	HMKCU94	97898 02/26/97 209044 05/15/97	pSport1	213	1079	502	896		739	515	1	23	24	43
32	HRDEW41	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	42	1983	1092	1983	27	27	344	1	11	12	520

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
32	HRDEW41	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	214	3791	2757	3357		2030	516	1			3
33	HTOJN06	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	43	1406	1	695		19	345	1	19	20	39
34	HBGDA21	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	44	1391	851	1153	74	74	346	1	30	31	234
34	HBGDA21	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	215	1334	822	1036		638	517	1	18	19	174
35	HFGAK75	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	45	1569	768	1569	14	14	347	1	19	20	169
35	HFGAK75	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	216	1511	770	1404	844	844	518	1	32	33	43

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
36	HHPBD40	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	46	1924	1	1681	62	62	348	1	19	20	43
37	HOVCL83	97898 02/26/97 209044 05/15/97	pSport1	47	475	252	396	141	141	349	1	37	38	78
38	HBCAY62	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	48	346	1	346	61	61	350	1	19	20	24
39	HBICM48	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	49	1366	882	1300	177	177	351	1	30	31	274
39	HBICM48	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	217	642	192	581		448	519	1			13
40	HLTCL35	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	50	1405	110	1404	61	61	352	1	30	31	47

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
40	HLTCL35	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	218	1241	1	1241	172	172	520	1	21	22	30
41	HLHCK50	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	51	504	207	485	222	222	353	1			3
42	HRSAN45	97899 02/26/97 209045 05/15/97	ZAP Express	52	777	1	214	113	113	354	1	24	25	52
43	HSNBB14	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	53	602	1	419	41	41	355	1	59	60	132
43	HSNBB14	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	219	1080	186	686	399	399	521	1	26	27	47
44	HMA BL38	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	54	1749	222	1749	166	166	356	1	30	31	204

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
44	HMA BL38	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	220	1258	149	1190	254	254	522	1	18	19	26
45	HSKDK47	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	55	1896	596	1614	650	650	357	1	33	34	47
46	HOSFH03	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	56	1753	555	1753	414	414	358	1	18	19	73
46	HOSFH03	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	221	1693	554	1693		526	523	1	25	26	58
47	HOGAV75	97899 02/26/97 209045 05/15/97	pCMV Sport 2.0	57	1220	690	1024	128	128	359	1	30	31	102
47	HOGAV75	97899 02/26/97 209045 05/15/97	pCMV Sport 2.0	222	1196	712	1163		1097	524	1			19

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
48	HFCAl74	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	58	1049	362	1049	335	335	360	1	33	34	48
49	HAGBI17	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	59	1776	854	1737	189	189	361	1	30	31	179
49	HAGBI17	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	223	1791	979	1791	1164	1164	525	1	18	19	40
50	HLFBC91	97899 02/26/97 209045 05/15/97	pBluescript SK-	60	443	1	443	164	164	362	1	21	22	25
51	HPRCA31	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	61	2888	1909	2888	90	90	363	1	30	31	224
51	HPRCA31	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	224	2517	1597	2517	1953	1953	526	1	18	19	57

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
52	HPRCE95	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	62	1851	1568	1736	139	139	364	1	30	31	349
52	HPRCE95	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	225	2424	299	2309		530	527	1	17	18	21
53	HHTLC66	97899 02/26/97 209045 05/15/97	ZAP Express	63	3542	883	3492	964	964	365	1	25	26	467
54	HMADJ02	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	64	883	237	883	229	229	366	1	30	31	152
54	HMADJ02	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	226	1080	242	1033	436	436	528	1	24	25	39
55	HPRCU93	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	65	1541	1	1541	236	236	367	1	30	31	373

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
55	HPRCU93	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	227	1336	4	1336	946	946	529	1	25	26	128
56	HSAXS65	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	66	732	41	698	163	163	368	1	18	19	83
56	HSAXS65	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	228	2043	1133	1756	1262	1262	530	1	20	21	82
57	HKTAG35	209011 04/28/97	Uni-ZAP XR	67	629	1	629	264	264	369	1			21
57	HMEFX42	97899 02/26/97 209045 05/15/97	Lambda ZAP II	229	540	25	536	227	227	531	1			20
58	HHFHN61	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	68	1751	375	1751	95	95	370	1	19	20	227
59	HCWEF90	97899 02/26/97 209045 05/15/97	ZAP Express	69	508	1	508	22	22	371	1	30	31	79

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
59	HCWEF90	97899 02/26/97 209045 05/15/97	ZAP Express	230	448	9	448		1	532	1	22	23	75
60	HHGCM20	97899 02/26/97 209045 05/15/97	Lambda ZAP II	70	245	1	245	93	93	372	1	1	2	51
61	HFRAU10	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	71	361	1	361	1	1	373	1	30	31	61
61	HFRAU10	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	231	407	1	407	210	210	533	1	17	18	60
62	HATDT67	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	72	713	8	713	169	169	374	1	30	31	40
62	HATDT67	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	232	830	190	580	329	329	534	1	28	29	39

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	5' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
63	HOUBG93	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	73	862	1	862	67	67	375	1	30	31	44
63	HOUBG93	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	233	932	138	905	287	287	535	1			2
64	HMWEX24	97900 02/26/97 209046 05/15/97	Uni-Zap XR	74	4602	4162	4525	730	730	376	1	30	31	203
64	HMWEX24	97900 02/26/97 209046 05/15/97	Uni-Zap XR	234	2786	2406	2739	2577	2577	536	1	22	23	36
65	HSGBA84	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	75	1255	1	1195	112	112	377	1	28	29	29
66	HTOCD52	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	76	475	1	475	13	13	378	1	30	31	136

Gene No.	cDNA Clone ID	ATCC Deposit No. Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
66	HTOCD52	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	235	458	1	458	26	26	537	1			14
67	HTGCP16	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	77	465	25	299	74	74	379	1	33	34	41
68	HKIXR69	97900 02/26/97 209046 05/15/97	pBluescript	78	1907	1627	1730	26	26	380	1	30	31	468
68	HKIXR69	97900 02/26/97 209046 05/15/97	pBluescript	236	591	1	444	251	251	538	1			18
69	HETG109	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	79	1168	136	1168	267	267	381	1	20	21	29
70	HOBNC61	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	80	1285	132	1285	292	292	382	1	27	28	29

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
71	HFFAH94	97900 02/26/97 209046 05/15/97	Lambda ZAP II	81	1290	768	1054	701	701	383	1	21	22	138
72	HBIAI95	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	82	684	1	684	119	119	384	1	30	31	74
73	HSQEL25	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	83	2024	1609	1953	200	200	385	1	30	31	521
73	HSQEL25	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	237	1286	391	959		1204	539	1	9	10	11
74	HEBEG68	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	84	931	14	537	85	85	386	1	25	26	137
75	HBIAI39	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	85	825	59	802	66	66	387	1	30	31	186

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
75	HBIAB39	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	238	734	1	734	1	1	540	1	37	38	108
75	HBIAB39	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	239	809	80	794		294	541	1	15	16	106
76	HTXDU73	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	86	1238	36	918	17	17	388	1			1
77	HOEAS24	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	87	1460	9	1458	166	166	389	1	53	54	299
77	HOEAS24	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	240	2201	841	2080	507	507	542	1	43	44	136
77	HOEAS24	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	241	1661	311	1520	390	390	543	1	35	36	424

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
78	HTEIY30	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	88	1395	567	1395	639	639	390	1	36	37	49
79	HSKNE46	97900 02/26/97 209046 05/15/97	pBluescript	89	1186	352	1186	540	540	391	1	49	50	61
79	HSKNE46	97900 02/26/97 209046 05/15/97	pBluescript	242	1146	329	1146	564	564	544	1	21	22	39
80	HPMFL27	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	90	1821	1203	1614	1503	1503	392	1	30	31	79
81	HMWDN32	97900 02/26/97 209046 05/15/97	Uni-Zap XR	91	862	253	862	359	359	393	1	32	33	36
82	HPRAX55	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	92	696	349	696	98	98	394	1	30	31	180

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
82	HPRAX55	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	243	1350	265	1230	348	348	545	1	32	33	58
83	HHFFW36	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	93	1886	1	1759	197	197	395	1			21
84	HE2PL77	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	94	1774	742	1772	785	785	396	1	21	22	60
85	HSDFV29	209076 05/22/97	Uni-ZAP XR	95	2503	1	1648	206	206	397	1	32	33	152
85	HCQAV53	97901 02/26/97 209047 05/15/97	Lambda ZAP II	244	1529	72	911	191	191	546	1	20	21	33
86	HTPEG42	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	96	2801	418	2801	234	234	398	1	30	31	480
86	HTPEG42	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	245	1537	1	1537	125	125	547	1	21	22	367

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
87	HLHDR57	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	97	1631	916	1631	1	1	399	1	1	2	423
88	HAUAV32	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	98	504	26	504	197	197	400	1	23	24	78
88	HAUAV32	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	246	506	1	499	183	183	548	1	32	33	77
89	HNEBI60	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	99	1416	145	1416	456	456	401	1	18	19	74
89	HNEBI60	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	247	1348	84	1348	363	363	549	1	21	22	47
90	HSHCJ16	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	100	2847	1	2847		2	402	1			20

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
91	HTSEL31	97901 02/26/97 209047 05/15/97	pBluescript	101	1394	608	1346	602	602	403	1	23	24	87
92	HAUBL57	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	102	794	1	794	518	518	404	1	30	31	92
92	HAUBL57	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	248	1766	42	1766	356	356	550	1	30	31	168
92	HAUBL57	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	249	2664	47	1708		147	551	1	18	19	124
93	HODAS59	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	103	1544	898	1531	975	975	405	1			21
94	HE6CT48	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	104	871	106	871	248	248	406	1	34	35	174

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
94	HE6CT48	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	250	865	97	865	258	258	552	1	19	20	177
95	HMDAA61	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	105	404	1	404	16	16	407	1	21	22	64
95	HMDAA61	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	251	2082	852	2074	829	829	553	1	22	23	72
96	HAQBK61	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	106	1542	506	1542	122	122	408	1	51	52	280
96	HAQBK61	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	252	1482	508	1482	633	633	554	1	15	16	45
96	HCUHB01	209215 08/21/97	ZAP Express	253	834	1	834	82	82	555	1	40	41	251
97	HAQBF73	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	107	2327	1528	2327	465	465	409	1	30	31	284

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
97	HAQBF73	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	254	1508	885	1508		988	556	1			19
98	HAQBT94	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	108	1062	157	1062	172	172	410	1	28	29	187
99	HETHE07	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	109	2539	275	2501	903	903	411	1	30	31	237
99	HETHE07	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	255	2514	592	2431	176	176	557	1	30	31	217
99	HETHE07	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	256	2357	465	2288		1151	558	1	12	13	82
100	HLQAB52	97901 02/26/97 209047 05/15/97	Lambda ZAP II	110	1751	969	1751	4	4	412	1	46	47	192

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
100	HLQAB52	97901 02/26/97 209047 05/15/97	Lambda ZAP II	257	689	218	655	314	314	559	1	18	19	95
100	HEONN58	209119 06/12/97	pSport1	258	2377	5	2377	25	25	560	1	28	29	54
101	HCRAM28	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	111	1117	1	1117		1	413	1	19	20	21
101	HIBEK16	209627 02/12/98	Other	259	1193	69	1135	242	242	561	1	24	25	108
102	HE2BG03	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	112	1313	128	1313	271	271	414	1	30	31	51
102	HE2BG03	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	260	1262	26	1262	35	35	562	1	35	36	50
103	HEBDJ82	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	113	1654	553	1654	709	709	415	1			32

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
104	HCUBC79	97901 02/26/97 209047 05/15/97	ZAP Express	114	1171	540	1171	337	337	416	1	30	31	163
104	HCUBC79	97901 02/26/97 209047 05/15/97	ZAP Express	261	1179	626	1161	335	335	563	1	30	31	253
104	HCUBC79	97901 02/26/97 209047 05/15/97	ZAP Express	262	1162	629	1131	942	942	564	1			18
105	HSVAF07	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	115	842	373	800	100	100	417	1	65	66	174
105	HSVAF07	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	263	735	290	735			565	1			
105	HSVAF07	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	264	783	416	783		413	566	1	33	34	73

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
106	HT3AM65	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	116	1640	187	1470	581	581	418	1	30	31	50
106	HT3AM65	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	265	1638	301	1405	119	119	567	1	30	31	263
106	HT3AM65	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	266	1455	148	1188	438	438	568	1	24	25	70
107	HE6DK18	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	117	952	418	906	499	499	419	1	28	29	120
108	HEBEK93	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	118	1256	21	1079	301	301	420	1	30	31	159
108	HEBEK93	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	267	1086	25	1050	227	227	569	1	23	24	34

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
109	HJPCM10	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	119	1143	171	1051	175	175	421	1	50	51	154
109	HJPCM10	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	268	1003	21	1003	115	115	570	1	34	35	104
109	HJPCM10	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	269	1234	174	1015	232	232	571	1	27	28	132
110	HSXBL78	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	120	1782	1	1720	138	138	422	1	32	33	204
111	HOEAW81	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	121	610	18	609	50	50	423	1	30	31	67
111	HOEAW81	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	270	574	1	566	337	337	572	1	27	28	32

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
112	HOEAP41	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	122	526	185	375	143	143	424	1	21	22	25
113	HEAAR60	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	123	2081	1179	1976	48	48	425	1	30	31	299
113	HEAAR60	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	271	1731	889	1626	886	886	573	1	18	19	28
114	HTXGS75	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	124	1717	764	1640	76	76	426	1			13
115	HOVBA03	97902 02/26/97 209048 05/15/97	pSport1	125	804	1	804	145	145	427	1	15	16	198
115	HOVBA03	97902 02/26/97 209048 05/15/97	pSport1	272	1320	77	637	280	280	574	1	22	23	40

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
116	HGBGK76	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	126	431	1	431	73	73	428	1	38	39	47
116	HGBGK76	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	273	515	1	515	43	43	575	1	20	21	30
117	HBMUW78	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	127	3752	3465	3752	748	748	429	1	30	31	370
117	HBMUW78	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	274	2995	2738	2995	2777	2777	576	1	18	19	29
118	HASAS24	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	128	1144	669	1144	896	896	430	1			30
119	HSIDN55	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	129	1830	1234	1830	1265	1265	431	1			24

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
120	HGBGZ64	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	130	1864	1505	1741	1578	1578	432	1	37	38	53
121	H6EBJ64	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	131	2041	1	1214	46	46	433	1	35	36	176
121	H6EBJ64	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	275	1990	8	1128	71	71	577	1	16	17	92
122	HOECP43	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	132	2012	853	1986	1127	1127	434	1	22	23	77
123	H2CBV31	97902 02/26/97 209048 05/15/97	pBluescript SK-	133	1669	670	1632	962	962	435	1	25	26	32
124	HPCAD23	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	134	1565	281	1565	274	274	436	1	25	26	30

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
125	HSPAG15	97902 02/26/97 209048 05/15/97	pSport1	135	2007	1101	2007	1124	1124	437	1	39	40	69
126	HELGH31	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	136	1291	1	1180	107	107	438	1			19
127	HUSHH48	97902 02/26/97 209048 05/15/97	Lambda ZAP II	137	1906	1	1906	184	184	439	1	30	31	43
127	HUSHH48	97902 02/26/97 209048 05/15/97	Lambda ZAP II	276	2436	572	2436	726	726	578	1	30	31	42
128	HL YAU95	97902 02/26/97 209048 05/15/97	pSport1	138	1935	1044	1794	1183	1183	440	1	18	19	33
129	HHSCV65	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	139	1446	572	1347	585	585	441	1	25	26	53

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
130	HTTAD57	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	140	1109	639	1109	676	676	442	1	24	25	64
131	HEBGA37	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	141	497	9	497	95	95	443	1			34
132	HEBFU93	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	142	269	1	269	1	1	444	1	30	31	89
132	HEBFU93	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	277	782	408	781		571	579	1	31	32	70
133	HSGSC60	97902 02/26/97 209048 05/15/97	Lambda ZAP II	143	1269	55	1262	55	55	445	1	25	26	350
134	HPMGD24	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	144	1944	97	1871	306	306	446	1	16	17	49

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
135	HPTVC60	97902 02/26/97 209048 05/15/97	pBluescript	145	1021	526	1021	74	74	447	1	30	31	278
135	HPTVC60	97902 02/26/97 209048 05/15/97	pBluescript	278	961	524	961	545	545	580	1	23	24	110
136	HSKNE18	97902 02/26/97 209048 05/15/97	pBluescript	146	1285	5	1285	116	116	448	1	30	31	199
136	HSKNE18	97902 02/26/97 209048 05/15/97	pBluescript	279	1228	9	1228	324	324	581	1	26	27	30
137	HMWIF35	97902 02/26/97 209048 05/15/97	Uni-Zap XR	147	1386	169	1272	165	165	449	1	30	31	258
137	HMWIF35	97902 02/26/97 209048 05/15/97	Uni-Zap XR	280	1327	169	1208	160	160	582	1	23	24	71

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
138	HMWGI25	97902 02/26/97 209048 05/15/97	Uni-Zap XR	148	2098	721	2044	784	784	450	1	18	19	87
139	HSKGF03	97902 02/26/97 209048 05/15/97	pBluescript	149	1847	1689	1847	241	241	451	1	33	34	315
139	HSKGF03	97902 02/26/97 209048 05/15/97	pBluescript	281	799	1	799		243	583	1	12	13	47
140	HMSKE75	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	150	1569	113	1517	417	417	452	1	21	22	52
141	HCMISH30	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	151	1540	538	1540	48	48	453	1	30	31	383
141	HCMISH30	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	282	2196	270	2196	294	294	584	1	32	33	39

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
142	HTWCB92	97902 02/26/97 209048 05/15/97	pSport1	152	1719	690	1575	6	6	454	1	52	53	186
143	HBMDM46	97902 02/26/97 209048 05/15/97	pBluescript	153	863	1	863	195	195	455	1	26	27	163
143	HBMDM46	97902 02/26/97 209048 05/15/97	pBluescript	283	1185	277	1166	621	621	585	1			19
144	HFAMG13	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	154	1101	1	512	40	40	456	1	21	22	46
145	HFXHL79	97903 02/26/97 209049 05/15/97	Lambda ZAP II	155	2031	669	2031	411	411	457	1	23	24	105
145	HFXHL79	97903 02/26/97 209049 05/15/97	Lambda ZAP II	284	1634	615	1485	878	878	586	1	20	21	23

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
146	HSNAK17	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	156	1981	1458	1809	1592	1592	458	1	23	24	70
146	HSNAK17	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	285	1795	1458	1749	1562	1562	587	1	33	34	69
147	HCFBC03	97903 02/26/97 209049 05/15/97	pSport1	157	915	45	912	22	22	459	1	22	23	155
147	HCFBC03	97903 02/26/97 209049 05/15/97	pSport1	286	858	46	858	224	224	588	1	30	31	77
147	HSJAP03	209139 07/03/97	Uni-ZAP XR	287	915	1	915	22	22	589	1	22	23	155
148	HSKGO26	97903 02/26/97 209049 05/15/97	pBluescript	158	2117	51	1422	32	32	460	1	23	24	332
149	HCQAV96	97903 02/26/97 209049 05/15/97	Lambda ZAP II	159	2395	1509	2382	1440	1440	461	1			5

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
150	HSHCC16	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	160	2120	1223	2108	1416	462	1			14
151	HTLEF62	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	161	900	482	900	46	463	1	30	31	285
151	HTLEF62	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	288	1517	783	1517	1062	590	1			24
152	HTLAD94	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	162	1003	1	1003	288	464	1	30	31	80
152	HTLAD94	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	289	3865	217	1195	281	591	1	16	17	38
153	HTSFQ12	97903 02/26/97 209049 05/15/97	pBluescript	163	2196	1607	2180	1611	465	1	30	31	47

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
154	HE6FL83	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	164	1945	271	1840	299	299	466	1	63	64	96
154	HE6FL83	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	290	1910	279	1818	355	355	592	1	39	40	69
155	HTXFJ55	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	165	2933	489	2871	258	258	467	1	30	31	399
155	HTXFJ55	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	291	3276	486	2838		525	593	1	45	46	308
156	HJPCJ76	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	166	2243	343	2221		341	468	1			1
157	HLTED27	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	167	1816	1130	1816	284	284	469	1	31	32	273

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
157	HLTED27	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	292	1695	1098	1548	1306	1306	594	1			22
158	HMKBA64	97903 02/26/97 209049 05/15/97	pSport1	168	945	1	787	208	208	470	1	18	19	192
159	HNFIP24	97903 02/26/97 209049 05/15/97	pBluescript	169	902	46	816	19	19	471	1	26	27	234
160	HCELB21	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	170	1883	798	1869	1001	1001	472	1	45	46	105
160	HCELB21	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	293	1501	438	1501	510	510	595	1			24
161	HAWBA28	97903 02/26/97 209049 05/15/97	pBluescript SK-	171	2100	1642	2100	1722	1722	473	1	23	24	32

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
162	HSAAS44	97903 02/26/97 209049 05/15/97	pBluescript SK-	172	1930	187	1930	65	65	474	1	30	31	571
162	HSAAS44	97903 02/26/97 209049 05/15/97	pBluescript SK-	294	2683	183	2683	431	431	596	1			24
163	HAFAL73	97903 02/26/97 209049 05/15/97	pBluescript SK-	173	1509	962	1451	122	122	475	1	30	31	312
163	HAFAL73	97903 02/26/97 209049 05/15/97	pBluescript SK-	295	1454	961	1420	976	976	597	1			1
164	HSAWF26	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	174	3173	2197	2972	51	51	476	1	21	22	329
164	HSAWF26	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	296	828	52	828	305	305	598	1			8

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
165	HEAAL31	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	175	991	374	970	60	60	477	1	24	25	178
165	HEAAL31	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	297	2416	1387	2413	1473	1473	599	1	18	19	25
166	HFKFX55	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	176	1290	499	1290		688	478	1	25	26	52
167	H2LAO11	97903 02/26/97 209049 05/15/97	pBluescript SK-	177	2290	1	2290	173	173	479	1	22	23	62
168	HPFDZ95	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	178	349	1	349	11	11	480	1	21	22	27
168	HPFDZ95	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	298	545	1	545	17	17	600	1	21	22	27

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
169	HPTTU11	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	179	1509	294	1352	92	92	481	1	30	31	339
169	HPTTU11	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	299	1530	385	1530	562	562	601	1	23	24	61
170	HCFAE79	97904 02/26/97 209050 05/15/97	pSport1	180	1316	985	1250	995	995	482	1	26	27	32
171	HTEDJ34	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	181	777	1	777	51	51	483	1	30	31	48
171	HTEDJ34	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	300	997	244	997	300	300	602	1	23	24	29
172	HODCW06	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	182	791	1	791	14	14	484	1	29	30	38

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
173	HFTAR26	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	183	1405	346	1405	575	575	485	1	20	21	61
174	H2MBF44	97904 02/26/97 209050 05/15/97	pBluescript SK-	184	1596	75	1596	131	131	486	1	24	25	346
174	H2MBF44	97904 02/26/97 209050 05/15/97	pBluescript SK-	301	2345	75	2345	233	233	603	1	56	57	69
175	HE8BI92	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	185	2293	355	2288	67	67	487	1	30	31	237
175	HE8BI92	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	302	2369	2	1946		60	604	1	9	10	24
176	HFTBR48	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	186	1212	462	1180	257	257	488	1	30	31	200

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
176	HFTBR48	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	303	1181	424	1149	663	663	605	1	23	24	35
177	HE9CM64	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	187	1605	770	1554	166	166	489	1	30	31	351
177	HE9CM64	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	304	1537	719	1515		787	606	1	43	44	130
178	HATAV51	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	188	1516	960	1516	8	8	490	1	30	31	265
178	HATAV51	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	305	1493	1	1261	54	54	607	1	18	19	23
179	HAQAF27	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	189	681	287	681		401	491	1			25

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
180	HCEEK08	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	190	1014	703	1014	360	360	492	1	30	31	159
180	HCEEK08	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	306	577	1	577	175	608	1				6
181	HAFU18	97904 02/26/97 209050 05/15/97	pBluescript SK-	191	2779	2207	2630	1153	493	1	30		31	279
181	HAFU18	97904 02/26/97 209050 05/15/97	pBluescript SK-	307	2860	163	2860	21	21	609	1	30	31	232
181	HAFU18	97904 02/26/97 209050 05/15/97	pBluescript SK-	308	876	275	876	302	302	610	1	32	33	34
182	HETBY74	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	192	1923	30	1923	45	45	494	1	33	34	193

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
183	HTOAF35	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	193	2346	1160	2286	178	178	495	1	30	31	205
183	HTOAF35	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	309	2025	840	2025	971	971	611	1	18	19	21
184	HCRBX32	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	194	3054	2004	3054	434	434	496	1	11	12	147
184	HCRBX32	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	310	3026	1966	3026		2131	612	1			9
185	HEGB80	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	195	907	152	907	297	297	497	1	30	31	64
185	HEGB80	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	311	712	67	712	107	107	613	1	18	19	29

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of ORF
186	HFAMH74	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	196	1290	84	809	225	225	498	1	30	31	94
186	HFAMH74	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	312	1289	785	1289	927	927	614	1	28	29	30

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below).

It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

- 5 The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, *Gene* 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources
- 10 using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

Signal Sequences

- Methods for predicting whether a protein has a signal sequence, as well as the
- 15 cleavage point for that sequence, are available. For instance, the method of McGeoch, *Virus Res.* 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, *Nucleic Acids Res.* 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1
- 20 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heinje, *supra*.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

- In the present case, the deduced amino acid sequence of the secreted polypeptide
- 25 was analyzed by a computer program called SignalP (Henrik Nielsen et al., *Protein Engineering* 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results
- 30 shown in Table 1.

- As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., +
- 35 or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely

uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

10 Polynucleotide and Polypeptide Variants

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

15 "Identity" per se has an art-recognized meaning and can be calculated using published techniques. (See, e.g.: (COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., ed., Oxford University Press, New York, (1988); BIOCOMPUTING: INFORMATICS AND GENOME PROJECTS, Smith, D.W., ed., Academic Press, New York, (1993); COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, (1994);
20 SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heinje, G., Academic Press, (1987); and SEQUENCE ANALYSIS PRIMER, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, (1991).) While there exists a number of methods to measure identity between two polynucleotide or polypeptide sequences, the
25 term "identity" is well known to skilled artisans. (Carillo, H., and Lipton, D., SIAM J Applied Math 48:1073 (1988).) Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in "Guide to Huge Computers," Martin J. Bishop, ed., Academic Press, San Diego, (1994), and Carillo, H., and Lipton, D., SIAM J Applied Math 48:1073 (1988).
30 Methods for aligning polynucleotides or polypeptides are codified in computer programs, including the GCG program package (Devereux, J., et al., Nucleic Acids Research (1984) 12(1):387 (1984)), BLASTP, BLASTN, FASTA (Atschul, S.F. et al., J. Molec. Biol. 215:403 (1990), Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711 (using the local homology algorithm of Smith
35 and Waterman, Advances in Applied Mathematics 2:482-489 (1981).)

When using any of the sequence alignment programs to determine whether a particular sequence is, for instance, 95% identical to a reference sequence, the parameters are set so that the percentage of identity is calculated over the full length of the reference polynucleotide and that gaps in identity of up to 5% of the total number of nucleotides in the reference polynucleotide are allowed.

A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. 6:237-245 (1990).) The term "sequence" includes nucleotide and amino acid sequences. In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB search of a DNA sequence to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, and Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, and Window Size=500 or query sequence length in nucleotide bases, whichever is shorter. Preferred parameters employed to calculate percent identity and similarity of an amino acid alignment are: Matrix=PAM 150, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty=0.05, and Window Size=500 or query sequence length in amino acid residues, whichever is shorter.

As an illustration, a polynucleotide having a nucleotide sequence of at least 95% "identity" to a sequence contained in SEQ ID NO:X or the cDNA contained in the deposited clone, means that the polynucleotide is identical to a sequence contained in SEQ ID NO:X or the cDNA except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the total length (not just within a given 100 nucleotide stretch). In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to SEQ ID NO:X or the deposited clone, up to 5% of the nucleotides in the sequence contained in SEQ ID NO:X or the cDNA can be deleted, inserted, or substituted with other nucleotides. These changes may occur anywhere throughout the polynucleotide.

Further embodiments of the present invention include polynucleotides having at least 85% identity, more preferably at least 90% identity, and most preferably at least 95%, 96%, 97%, 98% or 99% identity to a sequence contained in SEQ ID NO:X or the cDNA contained in the deposited clone. Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the polynucleotides having at least 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity

will encode a polypeptide identical to an amino acid sequence contained in SEQ ID NO:Y or the expressed protein produced by the deposited clone.

Similarly, by a polypeptide having an amino acid sequence having at least, for example, 95% "identity" to a reference polypeptide, is intended that the amino acid
5 sequence of the polypeptide is identical to the reference polypeptide except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the total length of the reference polypeptide. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be
10 deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the
15 reference sequence or in one or more contiguous groups within the reference sequence.

Further embodiments of the present invention include polypeptides having at least 80% identity, more preferably at least 85% identity, more preferably at least 90% identity, and most preferably at least 95%, 96%, 97%, 98% or 99% identity to an amino acid sequence contained in SEQ ID NO:Y or the expressed protein produced by
20 the deposited clone. Preferably, the above polypeptides should exhibit at least one biological activity of the protein.

In a preferred embodiment, polypeptides of the present invention include polypeptides having at least 90% similarity, more preferably at least 95% similarity, and still more preferably at least 96%, 97%, 98%, or 99% similarity to an amino acid
25 sequence contained in SEQ ID NO:Y or the expressed protein produced by the deposited clone.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or
30 activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in
35 the human mRNA to those preferred by a bacterial host such as *E. coli*).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an

organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level.

Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

5 Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988
10 (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

 Moreover, ample evidence demonstrates that variants often retain a biological
15 activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible
20 amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

25 Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form
30 are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

 Thus, the invention further includes polypeptide variants which show
35 substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make

phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

5 The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid
10 substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham
15 and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the
20 protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues
25 Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues,
30 where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino
35 acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

Polynucleotide and Polypeptide Fragments

10 In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in
15 length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

20 Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, and 701 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly
25 recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the
30 deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, and 161 to the end of the coding
35 region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about"

includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998-4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at
5 least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to
10 methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if
15 it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is
20 meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')₂ fragments) which are capable of specifically binding to protein. Fab and F(ab')₂ fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred,
25 as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

Fusion Proteins

30 Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular
35 locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

5 Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the
10 polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

 Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of
15 immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., *Nature* 331:84-86
20 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., *J. Biochem.* 270:3958-3964 (1995).)

 Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion
25 proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified,
30 would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., *J. Molecular Recognition* 8:52-58 (1995); K. Johanson et al., *J. Biol.*
35 *Chem.* 270:9459-9471 (1995).)

 Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In

preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the claimed invention.

Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS,

293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage

analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library) .) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the

present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute
5 biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags"
10 which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an
15 individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely
20 small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from
25 polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the
30 present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present
35 invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell. Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (125I, 121I), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99mTc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, 131I, 112In, 99mTc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20

millicuries of ^{99m}Tc . The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention could be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules

may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

Immune Activity

5 A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells
10 from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

15 A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic
20 cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency
25 (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood
30 coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks
35 (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from

inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitus, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

Hyperproliferative Disorders

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstrom's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

- Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, 5 Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., 10 Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiolitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), 15 pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.
- 20 Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergillois, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, 25 Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Meningococcal), Pasteurellaceae Infections (e.g., Actinobacillus, 30 Heamophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, 35 respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria,

Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect
5 any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis,
10 Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide
15 of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide
20 of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

25 A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteoarthritis, periodontal
30 disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vascular (including vascular endothelium), nervous, hematopoietic, and
35 skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stroke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

Chemotaxis

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotactic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotactic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotactic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat

disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

5 A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or
10 small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural
15 receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell
20 membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

25 The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations,
30 polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

35 Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The

antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity, and (b) determining if a biological activity of the polypeptide has been altered.

Other Activities

A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, circadian rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining

whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the

amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at
5 least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at
10 least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a
15 polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in
20 the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1;
25 and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining
30 whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of
35 polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an

amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1;
and a complete amino acid sequence of a protein encoded by a human cDNA clone
identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the
ATCC Deposit Number shown for said cDNA clone in Table 1.

5 Also preferred is the above method wherein said step of comparing sequences is
performed by comparing the amino acid sequence determined from a polypeptide
molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a
biological sample which method comprises a step of detecting polypeptide molecules in
10 said sample, if any, comprising an amino acid sequence that is at least 90% identical to
a sequence of at least 10 contiguous amino acids in a sequence selected from the group
consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as
defined in Table 1; and a complete amino acid sequence of a secreted protein encoded
by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained
15 in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type
of a biological sample, which method comprises a step of detecting polypeptide
molecules comprising an amino acid sequence in a panel of at least two amino acid
sequences, wherein at least one sequence in said panel is at least 90% identical to a
20 sequence of at least 10 contiguous amino acids in a sequence selected from the above
group.

Also preferred is a method for diagnosing in a subject a pathological condition
associated with abnormal structure or expression of a gene encoding a secreted protein
identified in Table 1, which method comprises a step of detecting in a biological sample
25 obtained from said subject polypeptide molecules comprising an amino acid sequence in
a panel of at least two amino acid sequences, wherein at least one sequence in said panel
is at least 90% identical to a sequence of at least 10 contiguous amino acids in a
sequence selected from the group consisting of: an amino acid sequence of SEQ ID
NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid
30 sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA
Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number
shown for said cDNA clone in Table 1.

In any of these methods, the step of detecting said polypeptide molecules
includes using an antibody.

35 Also preferred is an isolated nucleic acid molecule comprising a nucleotide
sequence which is at least 95% identical to a nucleotide sequence encoding a
polypeptide wherein said polypeptide comprises an amino acid sequence that is at least

90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated

polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

Examples

Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

	<u>Vector Used to Construct Library</u>	<u>Corresponding Deposited Plasmid</u>
	Lambda Zap	pBluescript (pBS)
	Uni-Zap XR	pBluescript (pBS)
20	Zap Express	pBK
	lafmid BA	plafmid BA
	pSport1	pSport1
	pCMVSPORT 2.0	pCMVSPORT 2.0
	pCMVSPORT 3.0	pCMVSPORT 3.0
25	pCR [®] 2.1	pCR [®] 2.1

Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Altting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Altting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polylinker to the T7 and T3 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation

of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.

Vectors pSport1, pCMVSPORT 2.0 and pCMVSPORT 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors
5 contain an ampicillin resistance gene and may be transformed into E. coli strain DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into E. coli strain XL-1 Blue. Vector pCR®2.1, which is available from Invitrogen, 1600 Faraday Avenue,
10 Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the
15 corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone
20 identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited
25 sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported.
30 The oligonucleotide is labeled, for instance, with ³²P-γ-ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as
35 those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection

agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 μ l of reaction mixture with 0.5 μ g of the above cDNA template. A convenient reaction mixture is 1.5-5 mM $MgCl_2$, 0.01% (w/v) gelatin, 20 μ M each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., *Nucleic Acids Res.* 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to

remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

10

Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide

A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

15

Example 3: Tissue Distribution of Polypeptide

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P³² using the rediprime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

20
25

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

30

Example 4: Chromosomal Mapping of the Polynucleotides

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This primer set is then used in a polymerase chain reaction under the following set of

35

conditions : 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on
5 either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

Example 5: Bacterial Expression of a Polypeptide

10 A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product
15 into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

20 The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r). Transformants are
25 identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The
30 cells are grown to an optical density 600 (O.D.⁶⁰⁰) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by
35 centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic

agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high
5 affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with
10 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in
15 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM imidazole. Imidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further includes
20 an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number XXXXXX.) This vector contains: 1) a neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence,
25 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (lacIq). The origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with NdeI and
30 XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible
35 enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

Example 6: Purification of a Polypeptide from an Inclusion Body

- 5 The following alternative method can be used to purify a polypeptide expressed in *E. coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

 Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at
10 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

- 15 The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

- 20 The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

- Following high speed centrifugation (30,000 xg) to remove insoluble particles,
25 the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

- To clarify the refolded polypeptide solution, a previously prepared tangential
30 filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a

stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem
5 columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0
10 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from
15 Commassie blue stained 16% SDS-PAGE gel when 5 μ g of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

Example 7: Cloning and Expression of a Polypeptide in a Baculovirus Expression System

20

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and
25 Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak *Drosophila* promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated
30 homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription,
35 translation, secretion and the like, including a signal peptide and an in-frame AUG as

required. Such vectors are described, for instance, in Luckow et al., *Virology* 170:31-39 (1989).

Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

Five µg of a plasmid containing the polynucleotide is co-transfected with 1.0 µg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., *Proc. Natl. Acad. Sci. USA* 84:7413-7417 (1987). One µg of BaculoGold™ virus DNA and 5 µg of the plasmid are mixed in a sterile well of a microtiter plate containing 50 µl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 µl Lipofectin plus 90 µl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a
5 "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the
10 suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the
15 recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 µCi of ³⁵S-methionine and 5 µCi ³⁵S-cysteine (available from Amersham) are added. The cells are
20 further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

25 **Example 8: Expression of a Polypeptide in Mammalian Cells**

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional
30 elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLV, HIV and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

35 Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden),

pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSPORT 2.0, and pCMVSPORT 3.0. Mammalian host cells that could be used include, human HeLa, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No. 209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the

naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("GeneClean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested
5 with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid
10 pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five μ g of the expression plasmid pC6 is cotransfected with 0.5 μ g of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that
15 confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri
20 dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 μ M, 2 μ M, 5 μ M, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 -
25 200 μ M. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

Example 9: Protein Fusions

The polypeptides of the present invention are preferably fused to other proteins.
30 These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394,827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time in vivo. Nuclear localization signals fused to the
35 polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having

more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in

5 Example 5.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

10 For example, if pC4 (Accession No.209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that
15 the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a
20 heterologous signal sequence. (See, e.g., WO 96/34891.)

Human IgG Fc region:

```
GGGATCCGGAGCCCAAATCTTCTGACAAAACTCACACATGCCCACCGTGCC
CAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCAAAACC
25 CAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTCACATGCGTGGTGGT
GGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACG
GCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAAC
AGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTG
AATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAACCCCC
30 ATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGT
GTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCT
GACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGTGGAGTGGGA
GAGCAATGGGCAGCCGGAGAACAACACTACAAGACCACGCCTCCCGTGCTGG
ACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGCA
35 GGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGC
ACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAGTGC
GACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)
```

Example 10: Production of an Antibody from a Polypeptide

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

10 In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., Nature 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell
15 Hybridomas, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at
20 about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line
25 (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (Gastroenterology 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

30 Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a
35 mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody

whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

5 It will be appreciated that Fab and F(ab')₂ and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')₂ fragments). Alternatively, secreted protein-binding fragments can be produced through the application of
10 recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art.
15 (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

20

Example 11: Production Of Secreted Protein For High-Throughput Screening Assays

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in
25 Examples 13-20.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well
30 (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at 2×10^5 cells/well in .5ml
35 DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in

5 Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of

10 transfections.

Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off

15 PBS rinse, and person B, using a 12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (see below) with 2mm glutamine and 1x penstrep.

20 (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B

25 adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

30 It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an

35 activity in a particular assay.

HGS-CHO-5 medium formulation:**Inorganic Salts**

CaCl ₂ (anhyd)	116.6 mg/L
CuSO ₄ ·5H ₂ O	0.00130
Fe(NO ₃) ₃ ·9H ₂ O	0.050
FeSO ₄ ·7H ₂ O	0.417
KCl	311.80
MgCl ₂	28.64
MgSO ₄	48.84
NaCl	6995.50
NaHCO ₃	2400.0
NaH ₂ PO ₄ ·H ₂ O	62.50
Na ₂ HPO ₄	71.02
ZnSO ₄ ·7H ₂ O	.4320

5 Lipids

Arachidonic Acid	.002 mg/L
Cholesterol	1.022
DL-alpha-Tocopherol-Acetate	.070
Linoleic Acid	0.0520
Linolenic Acid	0.010
Myristic Acid	0.010
Oleic Acid	0.010
Palmitric Acid	0.010
Palmitic Acid	0.010
Pluronic F-68	100
Stearic Acid	0.010
Tween 80	2.20

Carbon Source

D-Glucose	4551 mg/L
-----------	-----------

Amino Acids

L- Alanine	130.85 mg/ml
L-Arginine-HCL	147.50
L-Asparagine-H ₂ O	7.50
L-Aspartic Acid	6.65
L-Cystine-2HCL-H ₂ O	29.56
L-Cystine-2HCL	31.29
L-Glutamic Acid	7.35
L-Glutamine	365.0
Glycine	18.75
L-Histidine-HCL-	52.48

H ₂ O	
L-Isoleucine	106.97
L-Leucine	111.45
L-Lysine HCL	163.75
L-Methionine	32.34
L-Phenylalanine	68.48
L-Proline	40.0
L-Serine	26.25
L-Threonine	101.05
L-Tryptophan	19.22
L-Tyrosine-2Na-2H ₂ O	91.79
L-Valine	99.65

Vitamins

Biotin	0.0035 mg/L
D-Ca Pantothenate	3.24
Choline Chloride	11.78
Folic Acid	4.65
i-Inositol	15.60
Niacinamide	3.02
Pyridoxal HCL	3.00
Pyridoxine HCL	0.031
Riboflavin	0.319
Thiamine HCL	3.17
Thymidine	0.365
Vitamin B ₁₂	0.680

Other Components

HEPES Buffer	25 mM
Na Hypoxanthine	2.39 mg/L
Lipoic Acid	0.105
Sodium Putrescine-2HCL	0.081
Sodium Pyruvate	55.0
Sodium Selenite	0.0067
Ethanolamine	20uM
Ferric Citrate	0.122
Methyl-B-Cyclodextrin complexed with Linoleic Acid	41.70
Methyl-B-Cyclodextrin complexed with Oleic Acid	33.33
Methyl-B-Cyclodextrin complexed with Retinal Acetate	10

5

Adjust osmolarity to 327 mOsm

Example 12: Construction of GAS Reporter Construct

One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proximal region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

	<u>ISRE</u> <u>Ligand</u>	<u>JAKs</u>				<u>STATs</u>	<u>GAS(elements) or</u>
		<u>tyk2</u>	<u>Jak1</u>	<u>Jak2</u>	<u>Jak3</u>		
5	<u>IFN family</u>						
	IFN-a/B	+	+	-	-	1,2,3	ISRE
	IFN-g		+	+	-	1	GAS
	(IRF1>Lys6>IFP)						
	Il-10	+	?	?	-	1,3	
10	<u>gp130 family</u>						
	IL-6 (Pleiotrohic)	+	+	+	?	1,3	GAS
	(IRF1>Lys6>IFP)						
	Il-11(Pleiotrohic)	?	+	?	?	1,3	
15	OnM(Pleiotrohic)	?	+	+	?	1,3	
	LIF(Pleiotrohic)	?	+	+	?	1,3	
	CNTF(Pleiotrohic)	-/+	+	+	?	1,3	
	G-CSF(Pleiotrohic)	?	+	?	?	1,3	
	IL-12(Pleiotrohic)	+	-	+	+	1,3	
20	<u>g-C family</u>						
	IL-2 (lymphocytes)	-	+	-	+	1,3,5	GAS
	IL-4 (lymph/myeloid)	-	+	-	+	6	GAS (IRF1 = IFP
	>>Ly6)(IgH)						
25	IL-7 (lymphocytes)	-	+	-	+	5	GAS
	IL-9 (lymphocytes)	-	+	-	+	5	GAS
	IL-13 (lymphocyte)	-	+	?	?	6	GAS
	IL-15	?	+	?	+	5	GAS
30	<u>gp140 family</u>						
	IL-3 (myeloid)	-	-	+	-	5	GAS
	(IRF1>IFP>>Ly6)						
	IL-5 (myeloid)	-	-	+	-	5	GAS
	GM-CSF (myeloid)	-	-	+	-	5	GAS
35	<u>Growth hormone family</u>						
	GH	?	-	+	-	5	
	PRL	?	+/-	+	-	1,3,5	
	EPO	?	-	+	-	5	GAS(B-
40	CAS>IRF1=IFP>>Ly6)						
	<u>Receptor Tyrosine Kinases</u>						
	EGF	?	+	+	-	1,3	GAS (IRF1)
45	PDGF	?	+	+	-	1,3	
	CSF-1	?	+	+	-	1,3	GAS (not IRF1)

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:
5':GCGCCTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTCCCCG
10 AAATGATTTCCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:
5':CTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCGAAATG
15 ATTTCCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCC
CTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGC
CCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCTCGGC
CTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTT
20 TGCAAAAAGCTT:3' (SEQ ID NO:5)

25 With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not contain a

neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using
5 Sall and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

10 Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be
15 substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, IL-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

20 **Example 13: High-Throughput Screening Assay for T-cell Activity.**

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS
25 signal transduction pathway. The T-cell used in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-
30 SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml gentamicin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

35 Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI

+ 10% serum with 1% Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies) with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required
5 number of cells (10^7 per transfection), and resuspend in OPTI-MEM to a final concentration of 10^7 cells/ml. Then add 1ml of 1×10^7 cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Gentamicin, and 1% Pen-Strep. These cells are treated with supernatants
10 containing a polypeptide as produced by the protocol described in Example 11.

On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100
15 million cells) are required.

Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100, 000 cells per well).

After all the plates have been seeded, 50 ul of the supernatants are transferred
20 directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

The 96 well dishes containing Jurkat cells treated with supernatants are placed in
25 an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophane covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material
30 for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

Example 14: High-Throughput Screening Assay Identifying Myeloid Activity

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells.

- 5 Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

- 10 To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest 2×10^7 U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

- 15 Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$, 1 mM MgCl_2 , and 675 uM CaCl_2 . Incubate at 37°C for 45 min.

- 20 Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

- 25 These cells are tested by harvesting 1×10^8 cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of 5×10^5 cells/ml. Plate 200 ul cells per well in the 96-well plate (or 1×10^5 cells/well).

- 30 Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes,
5 EGR1 (early growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat pheochromocytoma cells) are known to proliferate and/or
10 differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor). The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

15 The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

20 5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)
5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the
25 EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and
30 allowed to air dry for 2 hr.

PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done
35 every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS (Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as 5×10^5 cells/ml.

Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to 1×10^5 cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ul of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

20

Example 16: High-Throughput Screening Assay for T-cell Activity

NF- κ B (Nuclear Factor κ B) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF- κ B regulates the expression of genes involved in immune cell activation, control of apoptosis (NF- κ B appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF- κ B is retained in the cytoplasm with I- κ B (Inhibitor κ B). However, upon stimulation, I- κ B is phosphorylated and degraded, causing NF- κ B to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF- κ B include IL-2, IL-6, GM-CSF, ICAM-1 and class I MHC.

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF- κ B promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF- κ B would be useful in treating diseases. For example, inhibitors of NF- κ B could be used to treat those diseases
5 related to the acute or chronic activation of NF- κ B, such as rheumatoid arthritis.

To construct a vector containing the NF- κ B promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF- κ B binding site (GGGGACTTTCCC) (SEQ ID NO:8), 18 bp of sequence complementary to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:
10 5':GCGGCCTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGAC
TTTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:

5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)
15 PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene)
Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:
20 5':CTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGACTTTCC
ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCA
TCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTTCTCCGCCCCATGGCTGACT
AATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTC
CAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTT:
25 3' (SEQ ID NO:10)

Next, replace the SV40 minimal promoter element present in the pSEAP2- promoter plasmid (Clontech) with this NF- κ B/SV40 fragment using XhoI and HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not
30 preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF- κ B/SV40/SEAP cassette is removed from the above NF- κ B/SEAP vector using restriction enzymes SalI and NotI, and inserted into a vector containing neomycin resistance. Particularly, the

NF- κ B/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and NotI.

Once NF- κ B/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described in Example 13. As a positive control, exogenous TNF alpha (0.1, 1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

Example 17: Assay for SEAP Activity

As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 μ l of 2.5x dilution buffer into Optiplates containing 35 μ l of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 μ l Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 μ l Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

Reaction Buffer Formulation:

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4

15	85	4.25
16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6
23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25
28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-3, used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO₂ incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

A stock solution of 1 mg/ml fluo-3 is made in 10% pluronic acid DMSO. To load the cells with fluo-3, 50 ul of 12 ug/ml fluo-3 is added to each well. The plate is incubated at 37°C in a CO₂ incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10⁶ cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-3 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10⁶ cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-3. The supernatant is added to the well, and a change in fluorescence is detected.

To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular

signaling even which has resulted in an increase in the intracellular Ca^{++} concentration.

Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity

5 The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase (RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

10 Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

20 Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

25 Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford, MA), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford, MA) are

used to cover the Loprodyn Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyn plates (20,000/200ml/well) and cultured overnight in complete medium.

- 5 Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na₃VO₄, 2 mM Na₄P₂O₇ and a cocktail of protease inhibitors (# 1836170) obtained from Boehringer Mannheim
- 10 (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation,
- 15 the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4°C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here.

- Generally, the tyrosine kinase activity of a supernatant is evaluated by
- 20 determining its ability to phosphorylate a tyrosine residue on a specific substrate (a biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

- 25 The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg₂₊ (5mM ATP/50mM MgCl₂), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl₂, 5 mM MnCl₂, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the
- 30 components gently and preincubate the reaction mix at 30°C for 2 min. Initiate the reaction by adding 10ul of the control enzyme or the filtered supernatant.

The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mM EDTA and place the reactions on ice.

- Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction
- 35 mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This

allows the streptavidin coated 96 well plate to associate with the biotinylated peptide. Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phosphotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as
5 above.

Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of
10 tyrosine kinase activity.

Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity

As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation
15 (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase,
20 Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then
25 rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1 and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C
30 until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts
35 filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. - As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and
5 Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

10

Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from
15 these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

20 PCR products is then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

25 PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals is identified by mutations not present in unaffected individuals.

30 Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenindeoxy-uridine 5'-triphosphate (Boehringer Mannheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the
35 corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv. et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect soluble polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on

the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale).
Interpolate the concentration of the polypeptide in the sample using the standard curve.

Example 23: Formulating a Polypeptide

5 The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for
10 purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 $\mu\text{g/kg/day}$ to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and
15 most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 $\mu\text{g/kg/hour}$ to about 50 $\mu\text{g/kg/hour}$, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes
20 and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal
25 patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

30 The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al.,
35 Biopolymers 22:547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D-(-)-3-hydroxybutyric

acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52,322; 5 EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

10 For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the 15 formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the 20 carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that 25 enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or 30 immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, 35 poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of

about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile.

5 Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized
10 formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or
15 more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the
20 present invention may be employed in conjunction with other therapeutic compounds.

Example 24: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or
normal expression level of a secreted protein in an individual can be treated by
25 administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

30 For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

Example 25: Method of Treating Increased Levels of the Polypeptide

Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

5 For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

10

Example 26: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and
15 separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS,
20 penicillin and streptomycin, is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

25 pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention can be amplified
30 using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions
35 appropriate for ligation of the two fragments. The ligation mixture is then used to

transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

5 The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

10 Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the
15 titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is being produced.

20 The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

25 The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference.

(1) GENERAL INFORMATION:

- 5 (i) APPLICANT: Human Genome Sciences, Inc. et al.
- (ii) TITLE OF INVENTION: 186 Human Secreted Proteins
- 10 (iii) NUMBER OF SEQUENCES: 644
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: Human Genome Sciences, Inc.
- 15 (B) STREET: 9410 Key West Avenue
- (C) CITY: Rockville
- (D) STATE: Maryland
- 20 (E) COUNTRY: USA
- (F) ZIP: 20850
- 25
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
- 30 (B) COMPUTER: HP Vectra 486/33
- (C) OPERATING SYSTEM: MSDOS version 6.2
- 35 (D) SOFTWARE: ASCII Text
- (vi) CURRENT APPLICATION DATA:
- 40 (A) APPLICATION NUMBER:
- (B) FILING DATE: March 6, 1998
- 45 (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
- 50 (A) APPLICATION NUMBER:
- (B) FILING DATE:
- 55

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: A. Anders Brookes, Esq.

(B) REGISTRATION NUMBER: 36,373

(C) REFERENCE/DOCKET NUMBER: PS002.PCT

(vi) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (301) 309-8504

(B) TELEFAX: (301) 309-8439

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 733 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGGATCCGGA	GCCCAAATCT	TCTGACAAAA	CTCACACATG	CCCACCGTGC	CCAGCACCTG	60
AATTCGAGGG	TGCACCGTCA	GTCTTCCTCT	TCCCCCAAAA	ACCCAAGGAC	ACCCTCATGA	120
TCTCCCGGAC	TCTGAGGTC	ACATGCGTGG	TGGTGGACGT	AAGCCACGAA	GACCCTGAGG	180
TCAAGTTCAA	CTGGTACGTG	GACGGCGTGG	AGGTGCATAA	TGCCAAGACA	AAGCCGCGGG	240
AGGAGCAGTA	CAACAGCACG	TACCGTGTGG	TCAGCGTCCT	CACCGTCCTG	CACCAGGACT	300
GGCTGAATGG	CAAGGAGTAC	AAGTGCAAGG	TCTCCAACAA	AGCCCTCCCA	ACCCCATCG	360
AGAAAACCAT	CTCCAAAGCC	AAAGGGCAGC	CCCAGAAACC	ACAGGTGTAC	ACCCTGCCCC	420
CATCCCGGGA	TGAGCTGACC	AAGAACCAGG	TCAGCCTGAC	CTGCCTGGTC	AAAGGCTTCT	480
ATCCAAGCGA	CATCGCCGTG	GAGTGGGAGA	GCAATGGGCA	GCCGGAGAAC	AACTACAAGA	540
CCACGCCTCC	CGTGCTGGAC	TCCGACGGCT	CCTTCTTCCT	CTACAGCAAG	CTCACCGTGG	600
ACAAGAGCAG	GTGGCAGCAG	GGGAACGTCT	TCTCATGCTC	CGTGATGCAT	GAGGCTCTGC	660
ACAACCACTA	CACGCAGAAG	AGCCTCTCCC	TGTCTCCGGG	TAAATGAGTG	CGACGGCCGC	720
GACTCTAGAG	GAT					733

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

10

Trp Ser Xaa Trp Ser
1 5

15

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 86 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GCGCCTCGAG ATTTCCTCGA AATCTAGATT TCCCGGAAAT GATTTCCTCG AAATGATTTC 60

30 CCGGAAATAT CTGCCATCTC AATTAG 86

35 (2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

45 GCGGCAAGCT TTTTGCAAAG CCTAGGC 27

50 (2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 271 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

60 CTCGAGATTT CCGGAAATC TAGATTTCCT CGAAATGATT TCCCGGAAAT GATTTCCTCG 60

AAATATCTGC CATCTCAATT AGTCAGCAAC CATAGTCCCG CCCCTAACTC CGCCCATCCC 120
GCCCCCTAACT CCGCCAGTT CCGCCATTC TCCGCCCCAT GGCTGACTAA TTTTITAT 180
5 TTATGCAGAG GCCGAGGCG CCTCGCCTC TGAGCTATTC CAGAAGTAGT GAGGAGGCTT 240
TTTGGAGGC CTAGGCTTTT GCAAAAAGCT T 271

10

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
15 (A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GCGCTCGAGG GATGACAGCG ATAGAACCCC GG 32

25

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
30 (A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GCGAAGCTTC GCGACTCCCC GGATCCGCCT C 31

40

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
45 (A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGGGACTTTC CC 12

55

(2) INFORMATION FOR SEQ ID NO: 9:

60 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 73 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GGGGCCTCGA GGGGACTTTC CCGGGGACTT TCCGGGGACT TTCCGGGACT TTCCATCCTG 60
10 CCATCTCAAT TAG 73

15 (2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 256 base pairs
(B) TYPE: nucleic acid
20 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

25 CTCGAGGGGA CTTTCCCGGG GACTTTCCGG GGACTTTCCG GGACTTTCCA TCTGCCATCT 60
CAATTAGTCA GCAACCATAG TCCCGCCCCCT AACTCCGCCC ATCCCGCCCC TAACTCCGCC 120
CAGTTCGGCC CATTCTCCGC CCCATGGCTG ACTAATTTTT TTTATTTATG CAGAGGCCGA 180
30 GGCCGCCTCG GCCTCTGAGC TATTCCAGAA GTAGTGAGGA GGCTTTTTTG GAGGCCTAGG 240
CTTTTGCAAA AAGCTT 256

35

(2) INFORMATION FOR SEQ ID NO: 11:

40 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 582 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GGCAGGAGGT AATTCTTACC AGAAATTTCC AGAGCATTAT GTAGGTAGAA AAAAATGCAA 60
50 GCAAGCTGTT AAAGATCTTG GATCCCATTA TATAGTATGT ATAGCTGAAA TCTGTAATTC 120
AATCACTTTT TCTCTTTTAT CCTCTAACCA AAAAATGTGT TAATTTTGCA TCCCAAATGT 180
TTTTAATCTT TGTATATTTT TTAAAAATCC TTTTCTCCTC ATCATGCTCT TTTTGTGGT 240
55 TGTAATAGA CTTACTTGCA CTTTGAAGAT GAGTTACTCC TTGTCATCTT ACAAATATGT 300
GATATGGTAA TTTTCATAAC AGATGTCAGT TTTGAACCAA GAATTGGTGA TTTGTTTATA 360
60 AGAAAAAATC TGGCTTCATT TCTGTGAAAT TGCTCTTTGA AAATTTCTTT TTACACGTGT 420

5 AAGCCAACTG AGATACCGTG ATGGTGTGA TTTCTTTCAA TGATGCTTAC CATCTATTTT 480
AGCCACTGAG CCTTTTATTA TTGTCTATT TGTAAGTTT ATTTGTCTTA ACTCATTTAA 540
TAAATATACT GTTTATCTGT TTCTGAAAAA AAAAAAAAAA AA 582

10

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

15

- (A) LENGTH: 465 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

20

60 GTTTGGGGGT GAGGCCGAGC TGCTGCGGG CTTCTCGCC GGCCAGGACA CAGCTACTCG
120 CACGGCGCG GCGCCTGGCT ATGATGTTCC TCACCCAGGG CGGGCCTCTG CCCTCTACTC
25 GTGCCAGGCC CACTTGCCAG GCAGGAGCCC TCCCCAAGCC TTCAGGGCTG CTCGGAGTCA 180
240 CCTGTGGAA TGGACTAAAA GGACCCTGT GTGGGAACAG GTGCTCCCCA AACACCCTGC
30 TGCTGGCTGC CAGGCAGGCC CTCTGGAAGG GAAGGGGAG GACTCATCAG GACCTCCCTG 300
360 GACCCCTGCA GGGCAGGCAG CTTGGGCCCC AGCCCAAGCA TTTGGCTCTG CTGCCCCCAA
420 GGGACAGGA AGCCTCTTGG GCCTCTTCCC TTCCTGGACA AGGCCCCCTG CCTTTCCTC
35 ACATAAACTG TACAGTATTT TCATTAAAAG CCTCTTTCAT AAAAA 465

40

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

45

- (A) LENGTH: 474 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

50

60 ATGCAATTCC TGCTCACAGC CTTCTGTTG GTGCCACTTC TGGCTCTTTG TGATGTCCCC
120 ATATCCCTAG GCTTCTCCCC CTCTAGAAG GGCTTCTGA TAGATTAGAA AATAAGAATG
55 AGTGACATTT CCTATGTGCA TATAAGAAGG AGCCACAAGA CATGTCTTTT AAATAAAAGG 180
240 ACAGTGCCA TCCTTTTAGC TGCCGAATAG AACCTGGTC TCATCCTCCT GGAGCTAGGC
300 CTTTAAACA GCTTCTGTGT TTCTCATTTG TCTCAGTGT TTGCCAGGGT TTTATCGGAA
60 AGATAATGTT CCGTTTAAAA TATTTCTTAA TGAGGCCGGG CGTGGTGGCT CACGCCTGTA 360

ACCCTAGCAM TTGGGGGCTG AGCGGGTGA TCACGAGGTC AGGAGATCGA GACCATCCTG 420

GSTAACATGG TGAAACCCCG TCTCTACTAA AAATACAAA AAAAAAAAAA AAAA 474

5

(2) INFORMATION FOR SEQ ID NO: 14:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 314 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

15

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

20 TTATGTTGGG GAGCAAGACC TGATAGCCAG CCTTTACATG GGAGTATAAT TCTGTCCTCC 60

ATCTCATAAG CCCAGTACC TGAGCCAGAA TGATTATAAC CAACCACACT GTCTCTTTAT 120

CATGGATGGC TTTAGCAGTA GGTTATTTTC ATCATTGCCA TTTGTAGCTC TACAGTGGTT 180

25 TATAGTAATT TCTCATCTTT TAAGTCTCTC CCTCAGTGCC TGTGTGTATC AAATCATTTG 240

CTCTCTCANG CAGTTGAGCT CTGCATCTC CCYTATGGGG GAGAGCTGTG TTGGAGAGAG 300

30 AGAATATNAC TTCC 314

(2) INFORMATION FOR SEQ ID NO: 15:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 613 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

40

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

45 CTCATATTGC CGTCTGGCTA AAAGTGAACA TGCCATTGAT CAATCTGCTT TTATTATATT 60

ATGTTCTCTAA TGGTGCAAG CAAGACAAGA AGTAGAAAGA AAGATGGTGT AAGCTCAAGA 120

ACCCACTAAA TCTATCCTAT GGCCTGGGTT CACCCAGCCT GCTTTGTGGA TTTTGTCTCA 180

50 CTATAACAGA GCTCCAAGG AGACTGCAGA GTCAGCTCCC TTAAGCACTG TAACTAAAGC 240

CTAACTCTTC CGTTCACCC AACAATGTYC CCAGCTCATC CTCTTTCCCR AAGTCCCCTT 300

55 TCTGCCCCAG ATGCGAATTG CATTTAACTA ATCCTCAAGT GAAATGTCCA CACAGRATTG 360

CATTTTAATT AGCATACCAT AGTTTTGTG CAAATTGCTT TTCAGARGAC TCCCATTGCA 420

GCTGCTCAGA GACGCTAAWG GCAGGCGCTC TTGAWGCTTT CCCGATAGCT TTCAGCTGCA 480

60 ATAGCTCTTA GGCAGAATGC CATGAGCGTC CTGCCCAACT GTATTACTGG GGAACACCTG 540

ATTGGCTAGA AGTTGATCCT CCTGTAAC TTCTGAGTTC TTTACATTTA CTCGTGAAAC 600
CCAAATATGC CAC 613

5

(2) INFORMATION FOR SEQ ID NO: 16:

10

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 356 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CCCCCCCCAT TGAACCTGG GCTGTGAAAG TTMTTGCCTG TGTGGGTCGT TCTGTGTGGC 60
GCCTGGTGTG TGGKTCCCAA CTCCTGTTCG AAAGTGGCAG CAGCCAATCA TGAAGCGCCC 120
TTATTTT TAG TTGCAGATGA CCAGTCTCC CCCCACAGC CTCTGTCTGG TCCCTCATTG 180
25 GTGAGTGGTC TGCCTGCCCA AGGAGCCTGA TTGGTGGGAA ATGGCATCAT CTAATATGAT 240
GGGAAGGCAT TTGGTCCTGG TTATGTTTAT TACAACATCA TTGCACTCTG GGACTCCAGT 300
CCCTGAAAC GTAATTTGTG GTGTTACCAA AGGACCACAG GGGAAAAAAA AAAAAA 356

30

(2) INFORMATION FOR SEQ ID NO: 17:

35

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 414 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

GAAACTANAT CCCGGGGCTT TTAACNGGTA CTGGGAAAT AAGTATTGGG TAATCACTAA 60
45 GNGACATTG ACTGCACCAA ACCAAAGCTA TAGAAAGAAA TGATTGACTT TTTAAATAT 120
ATTACATTA ACTGTCCTAG GATACTTCTC TTGAGGCTTT GGAAACTTC TTCCTTGAAA 180
50 TTTGCATATC CACTCCAGTT CTGTCACCAA AGATTTTAAT CTTGAGATCG CAATTTCTC 240
TCTCCAGAA AAAAGTACTA CAACAGGCTC AAGGGATATG CTTTGGTGGT CAAGGGATTA 300
CACTATGGTT TTCCTTCTGT TCACAATGGT ATTTACAGGA GACCTTGTCA TCAGAGGACG 360
55 TACTGAACTA TCTTTATGAC TTTGGATTG ATCAGAGGTT TAAAAAAA AAAA 414

60

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 469 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

10 AATCACCATT GCAATACAAA TGATCTGCCT GGTGAATGYT GAGCTGTACC CCACATTGCT 60
CAGGAACYTC GGAGTGATGG TGTGTTCCCTC CCTGTGTGAC ATAGGTGGGA TAATCACCCC 120
15 CTTCATAGTC TTCAGGCTGA GGGAGGTCTG GCAAGCCTTG CCCCTCATTT TGTTTGCGGT 180
GTGGGGCCTG CTTGCCCGG GAGTGACGCT ACTTCTTCCA GAGACCAAGG GGGTCGCTTT 240
GCCAGAGACC ATGAAGGACG CCGAGAACCT TGGGAGAAAA GCAAAGCCCA AAGAAAACAC 300
20 GATTTACCTT AAGGTCCAAA CTCAGAACC CTCGGGCACC TGAGAGAGAT GTTTTGCGGC 360
GATGTGCTGT TGGAGGGATG AAGATGGAGT TATCCTCTGC AGAAATTCCT AGACGCCTTC 420
25 ACTTCTCTGT ATTCTTCCTC ATACTGTCCT ACCCCCAAAT TAATATCAG 469

30 (2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 550 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

40 CCCCCCCCC CCCCACACT TTCAGGAGTC ACCCCCAGC ATTTGGGGTT GGGTTGGCCC 60
TACTCCAGCC TGGAGCTCCC TGAGGGAGCC TGCACTCCCT GCTCCCAATC CCCGCTACTG 120
GTGCAGGGAT GCAGCCTGGA GCTGGCGTCC TTGTTCTGGG CCTGCTGCTG CCGCCACCCC 180
45 AGAGCCCCAG CCTGTCTGA ATTGACATCA GTGCTTCCCT GAACTGCCTC CCCCACCCCT 240
GGGCATTATC CCAGGAACT TTATGTTTTT TAGAAGCTAA GCAGCTGCTG GGAATCAGGG 300
50 ACTGGTGAG GTAGGCTGAG TGGCAGCTCA GTCCTAGAAG GTCTCTGAAG ATCTGGACTG 360
AGGACCTTGC TACTCCCCAA GCCAGAGCCC ATCAGCCAGG CCTGCTGTGA GCCACCTGCC 420
TGTGGAGTGC TGAGCTCAAC CAAAGGCTGG CAAGCTCTGG GCCTCATTTA AGGGATTCTG 480
55 ATGAGCCGAT GGGCCCTGGA GGCAGCCCAT TAAAGCATCT GGCTCGTTTT TGGAAAAAAA 540
AAAAAAAAA 550

60

(2) INFORMATION FOR SEQ ID NO: 20:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 741 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
10 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

TCTTGAAGAG TGTACACTAC AGGATTATTA TAATGAAAGT TTATATCAAC AGGGTTTCGT 60
15 TGGCTCTGCA TATATTATAA GCAAAAGAGA TTGGTAAAGT GCCACAGTAT TCCAGATAAC 120
TTTTCAAGTG CGGCCTTTCT TCTCGTCTT TAATTTGAAA CCTAGATACA TGCAGTAAAA 180
20 ACTAGGAGAA TGACTTTTAC CCTTGGGGAC AGCCAAGTTT TGTTGATAAA CCTATTTTCT 240
AGCATGCCTT CAGGAAGTTG TGCCAGACCC TAGATTGTGA AGGACCCACT GTTCTTCTGT 300
TGTACGAGCT CCTGAACCA TTGTTTCTAG GACCAATGTC ACATCGCTTC ATGGGCATGG 360
25 NCCATGGGAG CATCTGGGTG ATAYCTGTCT ACAGTATTGG CTCTTCTGCG AGGCTGATAC 420
ACAAGGCCTC TCTTCCACAT GATCATTTGC AAACCTCCCC CAGCCCCTAC CATCCAATGT 480
30 GGAAGGAAAA CAAGAACTGC CTGAAGAAGA GTCCAAGCTA CAGATACACA GCGTGTGCAT 540
TGCGGCTGTC ACCTTCTCTC TCCCACTTCT GTATCCTCAG AGATGCTGCG TGGATGTTTC 600
CTTAACCTCA GCTGACTTCC CTGTGAATGT CTAATGCTAG TTCAGGGCCT CCAGGCATTG 660
35 ATTTGTACAG TGGTAACTCC CAATGAGGCT TCTGTTATCA TTTGGTGTGC TTTYTCTGTC 720
ATTAAAAGAA ATGATTTTCC C 741

40

(2) INFORMATION FOR SEQ ID NO: 21:

- 45 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 991 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GGCACGAGTC TCCCTGGGG AAGTTTTTCT TTTTCAGGAG GGAGGAGGGC TTTCCAGGT 60
AATGTGTCTA GAGTGTGGG CAGAAAATCT GGGACCACAC CACACCAGTT CTCTCCTTAA 120
55 TCCACGTCAT TTGCCTTCTA TCCAGCTAT GTTTCAGTG TCCTCTGGGT GTTCCAAGA 180
GCAACAAGAA ATGAATAAAT CTCTGGTGAG TTGTTTATTT GTTCTTCACT TTGTTTACA 240
60 CTGTATTTTC TGAGTTTATG GGTGCTCTGT AATTAAAAAG GAAAAGTAGA AATAAGTAAA 300

ACTCAGGTTG AAGGAAATAT ACATAAATAA GATAAAGCTG ACCTGTAGAT ATAGCAGGTT 360
ATAAAGCTTA GAGTTGTCTA AGTTGAGTGC AAATTTTCCT CTGATCTTTC TGATGCCGAA 420
5 CAAAAAGCA GTCATGTTTG TTATGTGATT GGAATGGAAC CCGAGAAGAG AGCATGCTGT 480
GTTCTGTGG GACAGGAAAG CTGCGTGCA CCAAGTCTGA ACCACCACCT TCATGGTGAC 540
10 ATAGATTATG TGCTGGAACA TATTTACAC CGGCCTGGCA GTAAACACTT GTAGTGTGT 600
GCAGTGAAA CGTCATCTT CCGCTAAAGC ACGGCGTGT GTGCAGCGGA AATGGTCATC 660
TGCTGCTAAA ACACAGCTTC CATCGTAATG TATGCTCCTT ACTCAAAGAG TGTGGTCCCA 720
15 AACAGCCTTT GGGAGGTCCT CTTGATTCA TGGATGAAAC CTGGAACATC TTGAGGACTG 780
AGTTAACCAT AGGTCCTTAA ATAACTCTCC ACACGTTTTT CTTAGTTTAT CTCTACATGC 840
20 AGGGTGTGCA GCAGCCTGTT CAAAGTCATA TTTTCTGGA AATATTTCCA GTGTTTATTT 900
GCACTTTAGC CCACTCTGTG TAGCCTTATT TCTTCTAAAC TCACCATTAA TCTGAATAAT 960
AGTCAAAATT AGGGGGACTG TATTTGCCTT A 991
25

(2) INFORMATION FOR SEQ ID NO: 22:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 653 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

35

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

CCACGCGTCC GGAATTCCTT TGAGGATCTT GGGCTATCTT TGACAGGGGA TTCTTGCAAG 60
40 TTGATGCTTT CTACAAGTGA ATATAGTCAG TCCCCAAAGA TGGAGAGCTT GAGTTCTCAC 120
AGAATTGATG AAGATGGAGA AAACACACAG ATTGAGGATA CGGAACCCAT GTCTCCAGTT 180
45 CTCAATTCTA AATTGTGTCC TGCTGAAAAT GATAGTATCC TGATGAATCC AGCACAGGAT 240
GGTGAAGTAC AACTGAGTCA GAATGATGAC AAAACAAAGG GAGATGATAC AGACACCAGG 300
GATGACATTA GTATTTTAGC CACTGGTTGC AAGGGCAGAG AAGAAACGGT AGCAGAAGAA 360
50 GTTGTATTG ATCTCACTTG TGATTCGGGG AGTCAGGCAG TTCCGTCACC AGCTACTCGA 420
TCTGAGGCAC TTTCTAGTGT GTTAGATCAG GAGGAAGCTA TGGAAATTAA AGAACCCAT 480
55 CCAGAGGAGG GGTCTTCAGG GTCTGAGGTG GAAGAAATCC CTGAGACACC TTGTGAAAGT 540
CAAGGAGAGG AACTCAAAGA AGAAAATATG GAGAGTGTTC CGTTGCACCT TTCTCTGACT 600
GAAACTCAGT CCCAAGGGTT GTGTCTTCGG AGGCATCCA AAAAAAAAAA AAA 653
60

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1486 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

5
10
15
20
25
30
35
40
45
50
55
60

GGCAGGCTGA CGACCTGCAA GCCACAGTGG CTGCCCTGTG CGTGCTGCCA GGTGGGGGAC 60
CCTGGGCAGG AAGCTGGCTG AGCCCCAAGA CCCCAGGGGC CATGGGCGGG GATCTGGTGC 120
TTGGCCTGGG GGCCTTGAGA CGCCGAAAGC GCTTGCTGGA GCAGGAGAAG TCTCTRGCCG 180
GCTGGGCACT GGTGCTGGCA SGARCTGGCA TTGGACTCAT GGTGCTGCAT GCAGAGATGC 240
TGTGGTTCGG GGGGTGCTCG GCTGTCAATG CCACTGGGCA CCTTTCAGAC AACTTTGGC 300
TGATCCCCAT CACATTCTTG ACCATCGGCT ATGGTGACGT GGTGCCGGGC ACCATGTGGG 360
GCAAGATCGT YTGCTGTGC ACTGGAGTCA TGGGTGTCTG CTGCACAGCC CTGCTGGTGG 420
CCGTGGTGGC CCGGAAGCTG GAGTTTAACA AGGCAGAGAA GCACGTGCAC AACTTCATGA 480
TGGATATCCA GTATACCAAA GAGATGAAGG AGTCCGCTGC CCGAGTGCTA CAAGAAGCCT 540
GGATGTTCTA CAAACATACT CGCAGGAAGG AGTCTCATGC TGCCCGCANG CATCAGCGCA 600
ANCTGCTGGC CGCCATCAAC GCGTTCCGCC AGGTGCCGCT GAAACACCGG AAGCTCCGGG 660
AACAAGTGAA CTCCATGGTG GACATCTCCA AGATGCACAT GATCCTGTAT GACCTGCAGC 720
AGAACTCTGAG CAGCTCACAC CGGGCCCTGG AGAAACAGAT TGACACGCTG GCGGGGAAGC 780
TGGATGCCCT GACTGAGCTG CTTAGCACTG CCTTGGGGCC GAGGCAGCTT CCAGAACCCA 840
GCCAGCAGTC CAAGTAGCTG GACCCACGAG GAGGAACCAG GCTACTTTCC CCAGTACTGA 900
GGTGGTGGAC ATCGTCTCTG CCACTCCTGA CCCAGCCCTG AACAAAGCAC CTCAAAGTGA 960
AGGACCAAAG GGGGCCCTGG CTTGGAGTGG GTTGGCTTGC TGATGGCTGC TGGAGGGGAC 1020
GCTGGCTAAA GTGGGKAGGC CTTGGCCAC CTGAGGCCCC AGGTGGGAAC ATGGTCACCC 1080
CCACTCTGCA TACCCTCATC AAAAACAATC TCACTATGCT GCTATGGACG ACCTCCAGCT 1140
CTCAGTTACA AGTGCAGGCG ACTGGAGGCA GGAATCCTGG GTCCCTGGGA AAGAGGGTAC 1200
TAGGGGCCCG GATCCAGGAT TCTGGGAGGC TTCAGTTACC GCTGGCCGAG CTGAAGAAGT 1260
GGGTATGAGG CTGGGCGGG GCTGGAGGTG GCGCCCCCTG GTGGGACAAC AAAGAGGACA 1320
CCATTTTTC AGAGCTGCAG AGAGCACCTG GTGGGGAGGA AGAAGTGTA CTCACCAGCC 1380
TCTGCTCTTA TCTTTGTAAT AAATGTTAAA GCCAGAAAAA AATAAAAAAA AAAAAAAAAA 1440

AACTCGAGGG GGGCCCRKAC CCAATCWCCC TATAGTAKAC GTANNN

1486

5

(2) INFORMATION FOR SEQ ID NO: 24:

- 10 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2323 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

CTTCGCCGTT TCTCCTGCCA GGGGAGGTCC CGGCTTCCCG TGGAGGCTCC GGACCAAGCC 60
CCTTCAGCTT CTCCCTCCGG ATCGATGTGC TGCCGCCGCC GCCGCCGCCG TCCCGCGTCC 120
20 TTCGGTCTCT GCTCCCGGA CCGGCTCCG CGCAGCCAGC CAGCATGTG GGGATCAAGA 180
AGCAAAAGAC GGAGAACCAG CAGAAATCCA CCAATGTAGT CTATCAGGCC CACCATGTGA 240
25 GCAGGAATAA GAGAGGGCAA GTGGTTGGAA CAAGGGGTGG GTTCCGAGGA TGTACCGTGT 300
GGCTAACAGG TCTCTCTGGT GCTGGGAAA ACAACGATAA GTTTTGCCCT GGAGGAGTAC 360
TTGTCTCCCA TGCCATCCCT GTTAATTCCT GGATGGGGAC AATGTCCGTC ATGGCCTTAA 420
30 CAGAATCCCC CAGATGGCTT CATGGCCCC AAAGCATGGA AGGTCCTGAC AGATTATTAC 480
AGGTCCCTGC AGAAGAACTA AGCCTTGGT CCAGAGTTC TTTCTGAAGT GCTCTTGAT 540
35 TACCTTTTCT ATTTTATGA TTAGATGCTT TGTATTAAAT TGCTTCTCAA TGATGCATTT 600
TAATCTTTTA TAATGAAGTA AAAGTTGTGT CTATAATTAA AAAAATATAT ATATATATAC 660
ACACACACAT ATACATACAA AGTCAAACTG AAGACCAAAT CTTAGCAGGT AAAAGCAATA 720
40 TTCTTATACA TTTCATAATA AAATTAGCTC TATGTATTTT CTA CTG CACC TGAGCAGGCA 780
GGTCCAGAT TTCTTAAGGC TTGTTTGAC CATGTGTCTA GTTACTTGCT GAAAAGTGAA 840
45 TATAITTTCC AGCATGTCTT GACAACCTGT ACTCTTCCAA TGTATTAT CAGTTGTAAA 900
ATATATCAGA TGTGCTCTT TCTGTACAAT TGACAAAAA AAAAATTTTT TTTTCTCACT 960
CTAAAAGAGG TGTGGCTCAC ATCAAGATTG TTCTTGATAT TTTACCTCAT GCTGTACAAA 1020
50 GCCTTAATGT TGTAAATCATA TCTTACGTGT TGAAGACCTG ACTGGAGAAA CAAAATGTGC 1080
AATAACGTGA ATTTTATCTT AGAGATCTGT GCAGCCTATT TCTGTCACAA AAGTTATATT 1140
55 GTCTAATAAG AGAAGTCTTA ATGGCCTCTG TGAATAATGT AACTCCAGTT ACACGGTGAC 1200
TTTTAATAGC ATACAGTGAT TTGATGAAAG GACGTCAAAC AATGTGGCGA TGTCGTGGAA 1260
AGTTATCTTT CCCGCTCTTT GCTGTGGTCA TTGTGCTTG CAGAAAGGAT GGCCCTGATG 1320
60

CAGCAGCAGC GCCAGCTGTA ATAAAAATA ATTCACACTA TCAGACTAGC AAGGCACTAG 1380
AACTGGAAAA GACCACAGAA AACAAAGAAT CCAACCCTTT CATCTTACAG GTGAACAAAC 1440
5 TGTGATGATG CACATGTATG TGTTTTGTAA GCTGTGAGCA CCGTAACAAA ATGTAAATTT 1500
GCCATTATTA GGAAGTGCTG GTGGCAGTGA AGAAGCACCC AGGCCACTTG ACTCCCAGTC 1560
TGGTGCCCTG TCTACACCAG ACAACACAGG AGCTGGGTCA GATTCCCTC AGCTGCTTAA 1620
10 CAAAGTTCCT CGAACAGAAA GTGCTTACAA AGCTGCCTTC TCGGATACTG AAAGGTCGAG 1680
TTTCTGAAC TGCAGTATT TTATTGCAGT TGAAAAA AAAAGCTAT TCCAAAGATT 1740
15 TCAAGCTGTT CTGAGACATC TTCTGATGGC TTTACTTCCT GAGAGGCAAT GTTTTACTT 1800
TATGCATAAT TCATTGTTGC CAAGGAATAA AGTGAAGAAA CAGCACCTTT TAATATATAG 1860
GTCTCTCTGG AAGAGACCTA AATTAGAAAG AGAAAAGTGT GACAATTTTC ATATTCTCAT 1920
20 TCTTAAAAAA CACTAATCTT AACTAACAAA AGTTCTTTTG AGAATAAGTT ACACACAATG 1980
GCCACAGCAG TTTGTCTTTA ATAGTATAGT GCCTATACTC ATGTAATCGG TTACTCACTA 2040
25 CTGCCTTTAA AAAAAAAAC CAGCATATTT ATTGAAAACA TGAGACAGGA TTATAGTGCC 2100
TTAACCGATA TATTTTGTGA CTTAAAAAT ACAITTTAAAA CTGCTCTTCT GCTCTAGTAC 2160
CATGCTTAGT GCAAATGATT ATTTCTATGT ACAACTGATG CTGTCTCTTA TTTTAATAAA 2220
30 TTTATCAGAG TGAAAAA AAAA AAAA AAAA AAAA AAAA AAAA 2280
AAAAA AAAA AAAA AAAA AAAA AAAA AAAA AAAA AAAA 2323
35

(2) INFORMATION FOR SEQ ID NO: 25:

- 40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 683 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

GGCAGGAGCC TGTGTGGTCA TGTTCCTCGT GGTGCAGTAC CTGACATGAG CCAGCCACGC 60
50 TCAGTGGCTG AACAGCATTC CCACAGCCTG CAAGTGTGTG TGTGTGTGAA AGAGAGAGGG 120
GGGCCCAGAG CCGCCTTTTG AAATGTTTGC CTGTCTGAAC TGTGAAGACA CTTGGGAGTG 180
ATGTGGTCT AATTTCCAAC CTGCTCTGTT TTCTGTGACA TCTTGGAGGG GAGCTAGTGC 240
55 CACACCATGC GCGGTGCTTA GAAATGAAAA AGTCCCGGT CTGTCTCTCT CACTCTCGCT 300
CTCATGGGG AGGGAAGAA TGGCTTTGGT GGCTTTGTTC ACACAGCTGA TGGTGTCTGG 360
60 GAAGGTGTCC ACAGTGAGCC TGTGTGCAGG ACTGTCCACA CGTTTCACAC TTGTACCAT 420

5 CAGGCCTTTC TGGTCCTGAT AGGGTGGAGC AAAAGTGGAA AGGAAAGGAA AGAGGCTTTT 480
CTCACAGCCA TTATATTAAA TAGTAGGTCG ATTCACATCT CGTGCTCCTG GCCACCTTCC 540
CCTGTGCCTC AGTGACATGT AGATGACTGA CTGCCAATAC TTGTCACCAT TCCCTGGAAG 600
CAGCTACCTA GGGGAAACAA GATGTAGTGC TATTGCCGAT AACAAAGTAAG ATTTTCCACA 660
10 CTAAAAA AAAA AAA 683

15 (2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 2036 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

25 CTGAGAAAGG AAAGCATTCG GATCTGCTGC AAAACACAT ATATCCATAA AGACTCATGT 60
TATTCAGAAA ACAGATTGTG AACACAATCA CATTCGCATG AATCCTTTAA AAGGAAGAAG 120
ACCTTAAAGT ATCTGCAAAT CTGAATTCTT ATTTATTCCT TCACTGAATA TAGAAACAAT 180
30 GGTATCTGA TTATTAGAGA TATTATTTTG GATATGTTAC TTATTAACCT GCTATGGCTG 240
GTAACCATGA TAAAGTCTGT TATTAATAAC AACATAATTC TTTTTTAAA GAAGAAAAGC 300
35 TTATTTTCA TTGACAGTGT ATAGATTAT CTACTTAGTT GTGTTTGCT ATTAGTGTTT 360
TAATTTTTTT TTTAAGTTGA GTGTTTGATA AATTTAAGA CCCTGTCCC ACCTGTMTT 420
GAGTCTGTG TTGACTACAG GTATATAGCY CAWITTAAAA ATCCTAAAGC AAAAGAATTT 480
40 TATTTATAAA AGAATCMAMC MGTTCATGC ATGAGGCTGT GAAGTCAGAT ATTTAGTAAT 540
AAAAGCAGCA GTGCCTTTTT TTGTATTAC CCAITGACCC CCACCAAATG CAACTGTTTT 600
45 ATATTAAGAA AATAGTAACA ATTTTAAAT CTCAGAGTAA AATCTATTC ACTACATGCT 660
TTTCCCCCT TGTCTGATT TAAGCAGTGT GACTTGGCA TCTCTACATT GTCCTAGGGA 720
CAGTGGTGTT CTACAATATT ATCATGTATG ATGTTTATT GTGCTTTT ATTATAGTG 780
50 GCTTCTTACC AGAAACAGTA GGAAGAAACA CATGAAGTGT GTACAAGACA TGAACATTG 840
CTGCTGATAT GTTGTTTTTT CACATGCTTT TGAGTTTCA CTTTTTAAAC GAGAGCCAGC 900
55 AAGCAAATA GATGTGGCTG GGTCTGCCTG TCCGGCGGC TYTTTGCACC GAGCTCTCAA 960
ATCCTGTGTA TTGAGGTTT CTTTGTGTA CTCAGGATTG GAGCTACAGC TGGCCCCCCC 1020
TCTCTCCCAT TCGTTTGAAG AGACACTGAG GGAAACAAGG GTTCTTTTG AGGTGTCCTT 1080
60

GGCTGCCTTT TACGGGATGG GAGCCTTCTC CGGATCTTTT GTTCTTCTGC ACCTCTTGTA 1140
GCTACTGCCG GTGCAAGGTT GTAGATGTTA TTCCCAGGA GCCTGGGCTK GGGGGCTGAG 1200
5 CTGGGCTGAA TGCAAAAGCA TGCAACCAGA AGGCGGGCAA GGGGAGGAAA AGCAGGCCTG 1260
GCCTCATTTGG TCCCCTGGAG ATGTCTGTAG CAGTCAGCTC CAGCTTGGGC CTGGGGAAGC 1320
AGCCTGACCA AGGCGCTCAG GTGTGCCTGT TACAAGAAGA ACCTGCAGAA GGATAATTTG 1380
10 CACATGGAGC TGTGATAACA CTAATGTTGA TTTTTTTTTT TTTTACAAGT CATCAGRGAT 1440
GTTTGCAAAG TGAGTTTAT TTTTTGTAA TTCCTTATC TTTACTTAA GGTGAATGTG 1500
15 TATTCCTCTG GGAGGAATAG GAAGAAAACA GGAATGTTAA TAATGTCGAA CAGAAAACCT 1560
CCTCCCTTAT TAATATATAA TCYTCATGTA TTTATGCCMT AATGTAAGCT GACTTTTAAA 1620
AAGCTTTCTT TTGTTGCATG CCCTGTGCAG GCATCTGTAT TGTACATGCA TGCCTTTCGT 1680
20 CCTGTTTTCC TGTATAAAGT TAGTGAACAA AGAAATATTT TTGCCCTAGT TCATGTTGCC 1740
AAGCAATGCA TATTTTTTAA ATTTGTCATA TATGGAAAGA GCATGTTTGT TACATGTAAA 1800
25 AGCTTTACTG ATATACAGAT ATACTAATGT TTGAAGATGC TGTTCCTTGC AAGTGTACAG 1860
TTTTCAAATG TTGTTACCAG TGAAACACCC TTGTGGTTTA AACTTGCTAC AATGTATTTA 1920
TTATTCATTT CCTCCCATGT AACTAAGAAT CATGGCTATA TTTTCATATCA ACGTTATATT 1980
30 GAAAGTGAAG GGAAATGATT AATACAAGGT TTTGTAACAA AAAAAANAA ANNAAA 2036

35

(2) INFORMATION FOR SEQ ID NO: 27:

40

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 717 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

GGCACGAGAT AACATAGGCA CAATAATACT GTATGTCTAC TTCTAGGATT ATAAGGAATT 60
AACATTGAGA TGACATTTCC ATTTGAGAAG AAAATAGTTG CTTTCAGTGC CTTTATTTTG 120
50 ATTCTCTGGAG AGAGCAGACT CGCACCAACA TTCAACCCCA GCGCTGATAT GACAGTAATC 180
CTCAGAGGCA GAGCCAGCA CAAAACAGCA ATGCTAGAAA GTTACAATTG GAAAGTTTCC 240
TGCCAGCTTC GGAATGACA CTGCAAAGCT GATGCCAGAA ACTGCCAGAG TAATTCTCCT 300
55 CATTAAGTCT CTACCCACCC ACTTTCAGCT CCCCAAATTA ACTAGTCAG TTGACTAATC 360
CTCTTTACCT TTATCATTTA GGTGAGGCAT TGCACAAAAA CTCTCGACTT TGCCATATAA 420
60 GGGCTGTGGT TCTCTGTGGT CCTGGATAAG AGGCATCACC ATTATCTGGA AACATGCAGT 480

AAATGCAGAT TCTTCATCTT CTCCCCAGAC CTCCTGAGTT AGAAATTCAC AAGTTCTCCA 540
GGTGATCTCA TACATGCTAA AGTTTGAGAA CCATTGAGTA AAGTTAATGC ATTAAGAAGA 600
GATTAGATAG GGATGGTGGC GTATCTTCCT ACAGTTTCCC TGTTAACAAG AAAGTCAGAG 660
GTCAGTTGAT CAGACATTAG ATTATTTATT GCTAAACTA AAAAAAATTA AAAAAA 717

10

(2) INFORMATION FOR SEQ ID NO: 28:

15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 495 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

GAATTCGGCA CGAGCAGCAT CCTAATTTTA GTTTGGAGAT GCATTCTAAA GGATCTTCTC 60
TATTGCTTTT TCTCCACAA TTAATCTTGA TTCTGCCTGT CTGTGCACAT TTGCATGAGG 120
AACTGAAC TGTTTTCAT AGGTAAATGA GAGACTGAGT TTTTTCATTT CTGAAGAGAA 180
AGGGCATT TG CTCCTACAAG CTGAAAGGCA CCCCTGGGTG GCTGGGGCCC TCGTGGGAGT 240
TTCTGGGGGA TTGACCCCTA CAACATGCAG TGGCCCTACA GAAAAACCTG CAACTAAAAA 300
TTATTTT TTA AAAAGGCTCC TCCAGGAAT GCATATAAGG GCTAATCACC CAGTATTTTG 360
ARGCTTCGAA GARGTAATAR AMCCCTGGAG AGAGAACTG AGACATGTAA GAGGGTGGGA 420
ATGACTCAGT GGTGGCACAC TATGGAGTCC TGCCCAACAG TAGCACACAT CAACCCACTA 480
CACAGAAATC CTAGG 495

40

(2) INFORMATION FOR SEQ ID NO: 29:

45

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 556 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

AGCTTAACGT CATGATTCAT TAGGGGAATG CAAGGCAAAA CCATGATGAG AATGCCCTA 60
GACACCTCTT AGAAGAGCTG CTAGAAAGGC AGACAGCACC AAGCGCTTAA ATGAGATGGG 120
GGCACTGGTG CTTCTTCTGT GCCTACTGGT AGGGGTGCAG CAGAGTGGTT CAGTCTGGGA 180
CAGTTAGCTG GACATCACGT GGACCAACA CACGCATTTC CTGGGTACT TACCAAGGAG 240

60

AATAGAAAGC AGGCAGATCT TTACAGCAGC TCTTACCTGW TTGCAAAACA ATGGAAATGC 300
5 CCACATGTCC ACAAACAAGT KTGTGGTCTG CCTGTGCCAT GAAGCACAGT GTGGCTGAGC 360
GTCAAGAGTC CCCACACTCA AAGGAGGCAG CAGATACAGG GCTGCACACT GTGTGATTCC 420
ACACATGTGA CATCTGGAC ACGGACATGC TGGATGGCAA AACGAGCATC GGGCTGAGAG 480
10 GACTGCTGAG AAGGGGAACG GGGCTGCTGG GATGTGGGTT GATTGTAGCA GTAGCTCATG 540
GAGATGTGAC CTCAAA 556

15

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:
20 (A) LENGTH: 434 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

CTAAATGGTG ACTGTGGCTT TGTCGAGACA GGGCCCAAAT GGTAGGTGTG AACACAACAT 60
GCACAGAATG AGGAGACATG CAGAGTGCTG AAATACTGTC CTGGACAGAT GTGTTACATG 120
30 ACTTTCCTTTT CAGCTTATTT CTGTGGCTG CCTTTGAAGA TAGAGCTTIG TTGATATTTA 180
CATTAACCA AATGTATATA YTAGTTCCA TTCTGACATG TTATTTAGCA AARGAAAAAR 240
35 GAGTAATTCT ACATCAGCAT CTTTAGTGCA TGCTAAAAGA TTAAAAATGT CTTTGGGGGA 300
ACATGTTTTG TATACATAAA TGTTAGATA GAAATATTTA TAGAATNCTC TATGTGAGTA 360
TTNATCTCCC TATGTATATT TATATCTAGA TGTGTCAATC TTTGTATTGA TATGAAATGC 420
40 TATGAATAGT GAGA 434

45

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:
50 (A) LENGTH: 715 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

CCACGCGTCC GATCTCACAG CTCGACACT ATTGCGAGCC ATACACAACC TGGTGTGAGG 60
AAACGTACTC CCAAATAAG CCCAAGATGC AAAGTTTGGT TCAATGGGGG TTAGACAGCT 120
60 ATGACTATCT CCAAATGCA CCTCCTGGAT TTTTCCGAG ACTTGGTGTT ATTGGTTTTG 180

CTGGCCTTAT TGGACTCCTT TTGGCTAGAG GTTCAAAAAT AAAGAAGCTA GTGTATCCGC 240
CTGGTTTCAT GGGATTAGCT GCCTCCCTCT ATTATCCACA ACAAGCCATC GTGTTTGCCC 300
5 AGGTCAGTGG GGAGAGATTA TATGACTGGG GTTACGAGG ATATATAGTC ATAGAAGATT 360
TGTGGAAGGA GAACTTTCAA AAGCCAGGAA ATGTGAAGAA TTCACCTGGA ACTAAGTAGA 420
10 AAACTCCATG CTCTGCCATC TTAATCAGTT ATAGGTAAAC ATTGGAATC CATAGAATAA 480
ATCAGTATTT CTACAGAAAA ATGGCATAGA AGTCAGTATT GAATGTATTA AATTGGCTTT 540
CTTCTTCAGG AAAAAGTAGA CCAGACCTCT GTTATCTTCT GTGAAATCAT CCTACAAGCA 600
15 AACTAACCTG GAATCCCTTC ACCTAGAGAT AATGTACAAG CCTAGAACT CCTCATCTC 660
ATGTTGCTAT TTATGTACCT AATTAAAACC CAAGTTAAAA AAAAAAAAAA AAAAA 715
20

(2) INFORMATION FOR SEQ ID NO: 32:

25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 486 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

GAGCCAGTGC CGGCGAAAGG GGACCTTCCT CTACTTCCTG CCACAGACCC TGTCCCCACA 60
35 CACTTCCTGC CCTGCTCTG CTGGGAGGCC ACTTCCTCCC CCAGTGCTGG ATTCCACCCC 120
CAGCTACCC TCAACATGG CCCCCTCTCT CCTCCTGCTT GCCCTCTCT GCTCCCTGGA 180
40 GGCTGTTCTG TCCTCCCTC TTGAAAAGCA ATGCCAGCTT CCTGGGATCT TCTGCCAACT 240
CCAGTACCA TGCCCTTTGC TCCTGTCAGC TCAGCTCCTC AAGGGAATTG TCTAMCCTCG 300
GTGTCTGCT TCCCTCCCTC AACCTCCTCA CCTGCTCCA AGCTGGCATC TGCCCCCTCA 360
45 CTGCACAGAA CGNTCCCCC ACCACCTGCC TTTACAGGGA GGAAGCAGCA ACATGGAAGA 420
ANCGAACTAT AGGGGCTACA ANGATGCTCA GCTCTGATCC CGAAGGCAA AAGNATCTTT 480
50 GGGCAC 486

(2) INFORMATION FOR SEQ ID NO: 33:

55 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 725 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
60 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

5 GTTCCTCTGG TAATAATTAG GTTATTCCCA GAAGCACAGT GTCATTCTTT AAATAAAAGC 60
TTTCCTGTTT AAAGCTTTTC AAAGGAGCAG ACCACCTTGA AGATTCCCC TAGGGTTGAT 120
ATGTGTCTAA TTCATTTTAT AAAAATTATT CTGTGCTTCA TTTTAAAGCT TTGGCTATAT 180
10 AGTCAGAAAT GTCCTAAATA ACAAATATT TTGTATTTAA TTTAGGGAAG ACTAAAGGGA 240
AGAAAAATGA AAACCTAGTC TTTATGTAAG CTCCAAGGAT ATTAGGGCTT AAAGGGCTTT 300
TCTAGTTTTA TGAGAAATTG TACTACTGAT TTTTATATAT TCCTGTTTTT GATGAACAGA 360
15 TCTCTGGGGA AATTGTTGAG TTACAATGGC ATTTCACGTG GATCCCTCTC AAGCTCAGAT 420
CAGTCTCTATA ACCCAATGAC AACCTGTCTC TTTGGTTTAC TGTCTGTGA AATGTCAGCT 480
20 CAAGTTTCCC AGAAGTCGTG TGTATTATGAT GAGTCAGAGT GCTTTTCTTC GGTGGGACAG 540
TTGCTGGCCC TCTTAATTTT GGTGTATGTG CTTCCAAGTA TCTAAACCTC CAGTCTGATC 600
TGTATATGCT ATCCTAACTG TTAATTGTAT TATTGATTAT GTTGATTATC TTGCTTGAAG 660
25 GTTCATACTT TTCAATTGTA TAGAAATAAA GTTTTTTTCT GCTTATAAAA AAAAAAAAAA 720
AAAAA 725

30

(2) INFORMATION FOR SEQ ID NO: 34:

- 35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 437 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
40 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

CACACAGCAT GCTGCCCTCA GACGTGTCCA TCCTGTACCA CATGAAAACG CTGCTGCTCC 60
45 TGCAAGATAC TGAGAGATTG AAGCATGCTC TGGAAATGTT CCCAGAACAT TGCACGATGC 120
CTCCTGCTTT TATTGGCTCT TGTGAAATC AAATTGGAAG ATCTTCAGTC CCAGCTGCAC 180
50 CCAAGTGA GAAGTATTCC AGGTCCATCC CCAAGGAACC AACACCGATG ACATGGACTC 240
AGGAATCTTA TAACCTACGT GGACTCTTTC CATCCGTACA TTGTCGTGCA CATGCCACTC 300
ATCACCTGGC GTGCCAGAT CCTCGCARG CAACACCTG TGATAATTCC AGGTGATTCT 360
55 CTACATCTGC AGCTTGAGGT TAGCCTCATA TCACATTACA TTCTACTAN AACNAAAAA 420
AAAAAAAAA AACTCNA 437

60

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 943 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

GGCACGAGCT GGAACAGAGA CTAATCCCA CGAACTGAC ATTGTTAAAC AACTAAAAC 60
AGAAGTACTT ACCTCTTGAA GATTTAATAT ATAATGGTTG ACATGATACA TGTACATGAT 120
15 GAATGACCAG ATGCTTATGG TCTACATTTT CCTTTATCCT GTTAGTATTA CCTTCCTTAA 180
TCTTTGTTC ATAACATGCT AATTCCTCTT CAGTGTTTAT TTTCTAGTGA CAGAATGCTA 240
20 ACATTTCCTA CACCCTGGCA GAAGGGAGAG AAATGTGTTT TGGGGTGGGT AACTAAATTT 300
TTGAGTGAAA TATCATAAGA TGANAATGGA AANAAGGAGA CACAAANAGT TATNACAAAA 360
AAACAATGGT TTTTITAGCC ATTTGACTGG CTCITTAAT AGTCTACAAG ACATTCACGT 420
25 TTAACATCAC TTTTAGTGAA ATAAAATGTG CCATACTAGT ATGTGCTTCA AAAGGGCAAA 480
TGTGCTTTAG TGCCCTAAGG CTAATTTTG GTCATTTGAC ATCAGAGATG TTGTAAGTAT 540
30 TGCACTTAAT ACGCACCTAT TTNTCAATAG TGTATTTTTT TGGNTAGCAT TTTTTTTACC 600
ACTATMTGT TGATAGCTTT TTGTTCTNTN AGGTTGNAAN ATGACAGTGC TNATNTCAAA 660
CAGATTACCC ATNTGCAGAA CTAAGGGAAG CNATTTATGT ATGAAAGNAA TTNTTGAATT 720
35 NGTCATTNTC AACCNITGNA TTAAAGCTTA GACTAAATAG TAATATATNG TGGGNAGGAT 780
TTTGGTTTTG TGATATTTNT GTGNATTAAG GNATAGATGT TAACCNITAT TTTGTAGNAA 840
40 AGTGANTTGT ATGTGGTTAA TTATAATAA AACTGGTACC AGGNAAAAAA AAAAAAAAAN 900
NAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAA 943

45

(2) INFORMATION FOR SEQ ID NO: 36:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 604 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

GGCACGAGAA ATCTTCATGC TGTAGTCACT CCAGACCATG GAGTGGCTTT CCAGCTGAAT 60
60 GAATCCTATG TCTCGCTGC AGGTGGTTGG TTTTCAATGT TCTTGCTAAT TTTTTTTCTA 120

TTGGATCTTG GGAGTTTCTT TTGTTTGCTC CTGTGTTTGC CCAGCTTTAA TAAAACCAGG 180
CGCAAACAAA AACCATAGCA TTCTGAACAA TAGGGGGCCC ACATTGGACC CAGTATGTCA 240
5 CTTTAATGGA CTTCAAGAAA AAATCTGAAT GGGAAAAATG ACACTAGGAA TGTATACTCC 300
ACACATTTTA TGCCATATAA TGGTGTGTTT TCTTAATTTT GTTTCTGTG GCGAAATGTG 360
10 GCTTTCAAAT TAAATGACC TTTTCTTCIT TGAAACTTTT TGTTTTGA CTGTATAATTA 420
AGGGTTTGA AAGATTCATA ATTCTGAGAG AGGTTTGCAA CCAGGAGATA CAAAGAAGTC 480
TCAGTAGTAA TCTGTTCAT GTGCTTTTAC AGCCAGCTAC ATTTAAGGAT GTATTAGTTA 540
15 CAGAAATTAT ATGTCTGTGT ATGTGTCTCT ACTCAATAAA GTACATGCCT CCACAAAAA 600
AAAA 604

20

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:
25 (A) LENGTH: 349 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

GTGAGTGCCC GGGAGCCCCG AGGCCCTGCC CCTAAGAAGG ATATCTYTRA CCGCTCCCTT 60
GTCCACACCC TAACCCCCCA GCTGCTCAGG CAGTGGGCAC ATGGCAGGGG CCTCACTGGG 120
35 GGCACATAGA GCATTTGGGG GACTGCGAGT GCTCACCTTT GACTTCCTGC AGGTCGGGGG 180
AAAACCAGAT CATGATGACC AAAGTYTACA TATTCTTGAT CTTTCATGGT CTGATCCTGC 240
40 CCTCCCTGGG TCTCACCAGG TATATGCCAC CACYTCTGY TCTAAATTCA GAATAAGAGT 300
CACATCAGGA GAGCACTGTC CCCAGGANAA TGCAAACGGG TTGGCAGCA 349

45

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:
50 (A) LENGTH: 672 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

GTAGTCGTTG CGGTTGCCGG GATGGCGAAG ATCTCGCCGT TTGAAGTCGT AAAACGCACC 60
TCGGTACCGG TGCTTGTTGG TTTGGTGATT GTWATCGTTG CTACAGAGCT GATGGTGCCA 120
60

GGAACGGCAG CAGCGGTCAC AGGCAAGTAA ATAGTAATGC CGGAGCAAGT TTCCTCCGGC 180
TTTATCATGT CACCCACTGT GGTATATGCG TTGTGGTCTG CCAACTTTGC CGTGAACAAT 240
5 TTCAGCAATA ATCAGATGGC GGCTGGCGCA ATATTCAAGA TAACGCCTGG CAGTGGTGCG 300
GCTGATGGTT CAGTGCCTGC GSCACCGTTT YTGCCGTATG TTGCACACCA GGNTCTTTAA 360
ACAGTTTTCG SACC CGGTTT AGCGTCAAGG GTTCAATGCC GGTCGGTAGC TCGTCCTTAG 420
10 GTTCACCGCG AGCATAAGCA TTAAACATCT CATCAATTTC CTTCTGGCTG GCGCTATCAA 480
TACTTTCCAG CATATGTTTA CGCTGGCGGA AACGGGTTAG CGTTTGCCCC ARCMGWTCAT 540
15 AGGCAATGGG CTTAATGAGA TAATCAAATA CACCACAACG TACGGCTTCA GACACCGTTT 600
CCATATCGCT GGCTGCAGTG GTAAACACCA CGTCGCCGGG ATAATGCCCC TGCACCAGTT 660
CATGCAGTAA AT 672
20

(2) INFORMATION FOR SEQ ID NO: 39:

25

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1908 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
30 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

AGAGTTGATA TTTT TAGAAA CAGTAATTTT ACTTTTAAGG AAATTGGCTA GCTCTTTGAC 60
35 TNNAGAGCTG TAGGAAGCTC AACATTTCTT TGTAGAGAAC GTTGCTTTTT TTGGATTGTA 120
CAGGTATAAA AACATTGCTT TTGTTGAATT GTATAGGTGT AAAAAGGGAA TAACTGTATG 180
40 CAGGTTTGAA AAGGAAATGT GCTTTAGGCA TGAGTCATAA GATGCCATG TACTTGTAGG 240
CATTTTATTT TCCTTTAGAA ATGGACATCA GCTCTTCTCT TCTGACTGGT AACACATAGC 300
CCCAAAGCAT GAGATTATTT TTCATTGGGT TTTTATTGTT GTTTAGTTT GGTTTGTTAC 360
45 GCCAGCCCAG TCTGCTGCG GAACACTGAC TCTGCTCTCT AATGAGAACA AAGTTAGAAA 420
TCTGCCGATA ACCTAAAATA ATTTAGAAAT GAATTAAAAA TGTGAAATCG GGTAAAGTG 480
50 ATGATGATAA AATAGCATGC AAGAAACAAG CTCCTTCCAT CAGACTTGGC TACTGTTTTC 540
TTCTGGTACG ATTTGGTTTG GAAGAGCCTC TTGTTTCCTT CTCTTTGGG TATGCTTTCG 600
TTTCTTAATA TGTTTGTAAC ATTATTGAGA TATAATTAC ATACCTTACA ATTCACTTAT 660
55 TTTAAGGTA CAATTTAGTG GTTTTAGTG TATTCACAAA GTTGTAAC CGTGACCACA 720
GTCAATTTTA GAACATTCG TTACCCCAA AAGAAACCCT GTACCCTGA GCAGTCACCT 780
60 CTCATTTTCT CCCAGTGCCC ACCCATCCC CGAGCCCCKG GAACCACTAA TCTATTTCTC 840

5 TCTCTGTAGA TTTGCTTATT CTGGTCATTT CATATAAATG GAATTCTACA ATATTGGGTC 900
TTTGGGACT GGCTTCCCAA ATATGATTTT CTATATGGAG TGAGAAAATT CTTCTCATCT 960
TGAGAACTCT TATTGCTGTG AAAGGGAGTG GTTGGTAAAA TCAATAGATT TCAGGCAAGA 1020
GGGCCAGATA CCTAACAGGT TTTCTCCGT GAATCTTATG CTGAGTAGTT TTTCTCATA 1080
10 ACCAAGCATT TATGATATAT TACTACTTAT AATACTGTGG CTAGTCTCTA GAATGGATGT 1140
TGAAATCTTT GCCTCCTCAG TCGGGAAGAG TCCTGCTAAA AATCAGGCTA AAAATCAGGC 1200
CAAAAATCAG GCCAAATGAC TTGGCAAATA ATTGACAAAG TGGTTTTCAC GTGTGTCTAT 1260
15 CTTTGCTAGC AGCTTGATA CCTCAGGCCA GGTGAGCTCC CCAAATTTCT TTTTTCATTT 1320
ACTCCAGTGA GTTCTGCTG TCTTTTCAA GTATGTACCA TAGGACTTAA AGGTGATTTG 1380
20 GATGCGTTGT AACACTGCTA AATATGCTAA GTACAGAATT TTATCTACAG TACTGTGAGA 1440
CAGTCAATTA TTGCCTAGGG TAGTTCAAAA ATATGATGTG AGCTAGTTAA GCCTTTGCTT 1500
GACTGATTTT AGTGATATTC AGAAGTGTGT ACCAATCAAG GCTCTTTAAA ATACGGAACG 1560
25 ACTCACTTAA TAACCAGGA ACCAGCCAAA TACTGTGCAG CCGCAGAATA TGCATATCAA 1620
TGAGTTGGAG GTGATTATTC TCTGTAATC CCTAATGATT GTTTTCTAAG CATTGTGGCT 1680
30 TCTCAGTGGC TTGACAGCAT CTTCTGGTT GTATGTGGCC TGTTTACATG ATGTATTGAA 1740
TAATGTGTGT TGTGTGAGC ATCAATGCCT GTAACACCAA ACTAAACACG TGTTTGTGGG 1800
ATATGTTTCC AATCTTTAAA TGACCTTGCC CTGTCCAATA AATAAATGAT TGTCTACCC 1860
35 TGTTAAAAA AAAAAAATT AAAAAACTG GNGGGGGGC CCGGTACN 1908

40

(2) INFORMATION FOR SEQ ID NO: 40:

45 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 458 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

CCTCAAAAA AAAAANGAAA GGAAAGAGGT CTCTACACAA GCCCGTGATT CTTCATGGCA 60
AGGGATAACA TCAGAAATGT TTCATTTYCK GCTATTAGTT TCCATTCCTT TCCCCATCCA 120
55 GGCATAAAGA GAAACAAAAG ACAATGATGG TATTCTCTGT GTCCTCAGCT TTGGCACTTT 180
TGTTGATGTT GCTAAGGAGC AGTGACCTTG CTA AAAAGAC TGAATAATCC ACCCACTGAA 240
TAGCTAACCT GGGGAGGAAA TGAAAATTTT CTTTGTGGAT CTCCCCAAAT CCATTGTTGT 300
60

CACCAGGCCC TCCCAGAACC TCCTCAGTTC CTTACAGTG CAACCCTGTG TACTTGGCCC 360
GCAACCCAAT AGTATTGTGC CTCATTTCAC CTTCCATGGG CAACTGCCCT CCCTTCTGGA 420
5 CATAAACCT CATATTTTAA ATNAAGTTGA AATTTGAA 458

10 (2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1153 base pairs
(B) TYPE: nucleic acid
15 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

20 GGCACAGAGC CTCGACCCA GGTGGTCTGG AGCCTGCCGG GAGAGTGGTG GCATCTGAGA 60
GGCTGGTCGT GGA CTGTGGT TGGGGGAGGT GGGAGCTGTT TTAACCGTGT GCCCCTCTC 120
CTGTGCCGGC GTGGGCATCC CCCGGGGCAG TGAACCCGG GCGCTCCTCC AGCTTCCGAG 180
25 TCCAGCCAGC CTGGGCGCGG GCGCGCCCC GAGACACCCG AGGAGTCCGT TCCTCCCTGG 240
TTACGTGGAC TGTGGAGCTG GTCTCTTGTG GCTCAGCGCC GTGCGGAGGT TGAAGCGTAC 300
30 CTGCGGAGGT CGCACCAGGG CGTGAGGAGG AGGAGGAAGG GCATGAGCCG AGCTTGAGGA 360
ATCCGTGCTC CAAACTCTAC ACTCAAGGAT GCACTGCGCA ACTCTGGTGG CGATGGGCTG 420
GGGCAGATGT CCTTGGAGTT CTACCAGAAG AAGAAGTCTC GCTGGCCATT CTCAGACGAG 480
35 TGCATCCCAT GGAAGTGTG GACGGTCAAG GTGCATGTGG TAGCCCTGGC CACGGAGCAG 540
GAGCGGCAGA TCTGCCGGA GAAGTGGGT GAGAACTCT GCGAGAAGAT CATCAACATC 600
40 GTGGAGGTGA TGAATCGGCA TGAGTACTTG CCCAAGATGC CCACACAGTC GGAGGTGGAT 660
AACGTGTTTG ACACAGGCTT GCGGGACGTG CAGCCCTACC TGTACAAGAT CTCCTTCCAG 720
ATCACTGATG CCTGGGCAC CTCAGTCACC ACCACCATGC GCAGGCTCAT CAAAGACACC 780
45 CTGCCCTCTG AGCGTCGCTG GATCTCTGGG AGCTCCTTGA TGGCTCCCAG ACCTTGGCTT 840
TTGGGAATTG CACTTTTGGG CCTTTGGGCT CTGGAACCTG CTCTGGGTCA TTGGTGAGAC 900
50 TTGGAAGGGG CAGCCCCCGC TGGCTTCTTG GTTTTGTGGT TGCCAGCCTC AGGTCATCCT 960
TTTAATCTTT GCTGACGTT CAGTCCTGCC TCTACTGTCT CTCCATAGCC CTGGTGGGGT 1020
CCCCCTCTT TCTCCACTGT ACAGAAGAGC CACCACTGGG ATGGGGAATA AAGTTGAGAA 1080
55 CATGAGTTTG GGCTGAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1140
AAAAAAAAA AAA 1153

60

(2) INFORMATION FOR SEQ ID NO: 42:

5

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1983 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

	GGCACGAGAG GGGCCGAGCC GACAAGATGT TCTTGCTGCC TCTTCCGGCT GCGGGGCGAG	60
15	TAGTCGTCCG ACGTCTGGCC GTGAGACGTT TCGGGAGCCG GAGTCTCTCC ACCGCAGACA	120
	TGACGAAGGG CCTTGTTTGA GGAATCTATT CCAAAGAAAA AGAAGATGAT GTGCCACAGT	180
	TCACAAGTGC AGGAGAGAAT TTTGATAAAT TGTTAGCTGG AAAGCTGAGA GAGACTTTGA	240
20	ACATATCTGG ACCACCTCTG AAGGCAGGGA AGACTCGAAC CTTTATGGT CTGCATCAGG	300
	ACTTCCCCAG CGTGGTGCTA GTTGGCCTCG GCAAAAAGGC AGCTGGAATC GACGAACAGG	360
25	AAAAGTGGCA TGAAGGCAAA GAAAACATCA GAGCTGCTGT TGCAGCGGGG TGCAGGCAGA	420
	TTCAAGACCT GGAGCTCTCG TCTGTGGARG TGGATCCCTG TGGAGACGCT CAGGCTGCTG	480
	CGGAGGGAGC GGTGCTTGGT CTCTATGAAT ACGATGACCT AAAGCAAAAA AAGAAGATGG	540
30	CTGTGTCGGC AAAGCTCTAT GGAAGTGGGG ATCAGGAGGC CTGGCAGAAA GGAGTCTCTG	600
	TTGCTTCTGG GCAGAACTTG GCACGCCAAT TGATGGAGAC GCCAGCCAAT GAGATGACGC	660
35	CAACCAGATT TGCCGAAATT ATTGAGAAGA ATCTCAAAAG TGCTAGTAGT AAAACCGAGG	720
	TCCATATCAG ACCCAAGTCT TGGATTGAGG AACAGGCAAT GGGATCATTC CTCAGTGTGG	780
	CCAAAGGATC TGACGAGCCC CCAGTCTTCT TGGAAATTC A CTACAAAGC AGCCCCAATG	840
40	CAAACGAACC ACCCCTGGTG TTTGTTGGGA AAGGAATTAC CTTTGACAGT GGTGGTATCT	900
	CCATCAAGGC TTCTGCAAAAT ATGGACCTCA TGAGGGCTGA CATGGGAGGA GCTGCAACTA	960
45	TATGCTCAGC CATCGTGTCT GCTGCAAAGC TTAATTTGCC CATTAAATATT ATAGGTCTGG	1020
	CCCCCTCTTTG TGAAAATATG CCCAGCGGCA AGGCCAACAA GCCGGGGGAT GTTGTTAGAG	1080
	CCAAAAACGG GAAGACCATC CAGGTTGATA AACTGATGC TGAGGGGAGG CTCATACTGG	1140
50	CTGATGCGCT CTGTTACGCA CACACGTTTA ACCCGAAGNT CATCCTCAAT GCCGCCACCT	1200
	TAACAGGTGC CATGGATGTA GCTTTGGGAT CAGGTGCCAC TGGGGTCTTT ACCAATTCAT	1260
55	CCTGGCTCTG GAACAAATC TTCGAGGCCA GCATTGAAAC AGGGGACCGT GTCTGGAGGA	1320
	TGCCTCTCTT CGAACATTAT ACAAGACAGG TTGTAGATG CCAGCTTGCT GATGTTAACA	1380
60	ACATTGGAAT ATACAGATCT GCAGGAGCAT GTACAGCTGC AGCATTCCTG AAAGAATTCTG	1440

5 TAACTCATCC TAAGTGGGCA CATTTAGACA TAGCAGGCGT GATGACCAAC AAAGATGAAG 1500
TTCCCTATCT ACGGAAAGGC ATGACTGGGA GGCCCAACAG GACTCTCATT GAGTTCTTAC 1560
TTCGTTTCAG TCAAGACAAT GCTTAGTTCA GATACTCAAA AATGTCTTCA CTCTGTCTTA 1620
AATGGACAG TTGAACATAA AAGGTTTTTG AATAAATGGA TGAAAATCTT TTAACGGAGA 1680
10 CAAAGGATGG TATTTAAAAA TGTAACAC AATGAAATTT GTATGCCTTG ATTTTTTTTTT 1740
CATTTACAC AAAGATTAT AAAGTAAAG TTAATATCTT ACTTGATAAG GATTTTAAAG 1800
ATACTCTATA AATGATTAAA ATTTTATAGAA CTTCCTAATC ACTTTTCAGA GTATATGTTT 1860
15 TTCATTGAGA AGCAAAATTG TAACTCAGAT TTGTGATGCT AGGAACATGA GCAAACTGAA 1920
AATTACTATG CACTTGTCAG AAACAATAA TGCAACTTGT TGTGCAAAAA AAAAAAAAAA 1980
20 AAA 1983

(2) INFORMATION FOR SEQ ID NO: 43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1406 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

35 ATGATGATGA CTTTGAAGAC GATTTTATTC CTCCTCTCC AGCTAAGCGC CTTGAGGTTA 60
ATAGTTGGAA AAGACTCTAT AGATATTGAC ATTTCTTCAA GGAGAAGAGA AGATCAGTCT 120
TTAAGGCTTA ATGCTAAGC NCTTGGTCTT AACTTGACCT GGGATAACTA CTTTAAAGAA 180
40 ATAAAAAATT CCAGTCAATT ATTCTCAAC TGAAAGTTTA GTGGCAGCAC TTCTATTGTC 240
CCTTCACTTA TCAGCATACT ATTGTAGAAA GTGTACAGCA TACTGACTCA ATTCTTAAGT 300
CTGATTTGTG CAAATTTTFA TCGTACTTTT TAAATAGCCT TCTTACGTGC AATTCTGAGT 360
45 TAGAGGTAAA GCCCTGTGT AAAATAAAGG CTCAAGCAA ATTGTACAGT GATAGCAACT 420
TTCCACACAG GACGTTGAAA ACAGTAATGT GGCTACACAG TTTTTTTAAAC TGTAAAGACA 480
50 TCAGCTGGCT CTTTAATATA TGAATAACA ATAATTTAAA ACAAATCATA GTAGCAGCAT 540
ATTAAGGGTT TCTAGTATGC TAATATCACC AGCAATGATC TTTGGCTTTT TGAATTTATTT 600
GCTAGATGTT TCCCCCTTGG AGTTTGTGCA GTTTCACACT GTTGTCTGGC CCAGGTGTAC 660
55 TGTGTGTGGC CTTTGTAAAT ATCGCAAACC ATTGGTGGG AGTCAGATTG GTTCTTAAA 720
AAAAAAAAA AAAACGACAT ACGTGACAGC TCACTTTTCA GTTCATTATA TGTACCGAGG 780
60 GTAGCAGTGT GTGGGATGAG GTTCGATACA GNCGTATTTA TTGCTTGTC TGTAAATTAA 840

AAACCTTGTA TTAACTCTT TTCAATCCTT TTAGATAAAA TTGTTCTTTG CAAGAATGAT 900
TGGTGCTTAT TTTTCAAAA ATTTGCTGTG AACAACTGA TGACAACAAG CAACATTAT 960
5 CTAATGAAC ACAGCTATCT TAATTTGGTT CTCAAGTTT TCTGKTGCAC TTGTAAAATG 1020
CTACAAGGAA TATTAAAAA ATCTATTAC TTTAACTTAT AATAGTTTAT GAAATAAAAA 1080
10 CATGAGTCAC AGCTTTTGT CTGTGGTAAC CTATAAAAA AGTTTGTCTT TGAGATTCAA 1140
TGTAAGAAC TGAAACAAT GTATATGTTG TAAATATTG TGTGTTGTA GAAATTTTGT 1200
TCATAAGAAA TTAAAGAAC TTACCAGGAA GGTTTTAAAG TTAGAAATAT TCCATGCCAA 1260
15 TAAATAGGA AATTATAAT ATATAGTTT AAGCCTGCAT CAGTGGGAGT CTGGGCTATG 1320
TAGTTATGTA GTTATTATGN AACCACCAAG ATTTTTTGG CTATTTACCG TAACCAAAGG 1380
20 GGCCGATTAA NTGGTTTGAA GNCTTG 1406

25 (2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1391 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

35 GGGCTGAAG GCGCRCGCC AGTCCGAGC AGTGCTCGCT CCTGCTCGGG GCGCTGCGGC 60
CCCGGGCGTC GCCATGACCA GTGAGCTGGA CATCTCGTG GGAACACGA CCCTTATCGA 120
CGAGGACGTG TATCGCCTCT GGCTCGATGG TTACTCGGTG ACCGACGCG TGGCCCTGCG 180
40 GGTGCGCTCG GGAATCCTGG AGCAGACTGG CGCCACGGCA GCGGTGCTGC AGAGCGACAC 240
CATGGACCAT TACCGCACCT TCCACATGCT CGAGCGGCTG CTGCATGCGC CGCCCAAGCT 300
45 ACTGCACCAG CTCATCTTCC AGATTCCGCC CTCCCGGAG GCACTACTCA TCGAGAGGTA 360
CTATGCCCTT GATGAGGCCT TTGTTGCGGA GGTGCTGGGC AAGAAGCTGT CCAAAGGCAC 420
CAAGAAAGAC CTGGATGACA TCAGACCAA AACAGGCATC ACCCTCAAGA GCTGCCGGAG 480
50 ACAGTTTGAC AACTTTAAAC GGTCTTCAA GGTGGTAGAG GAAATGCGG GCTCCCTGGT 540
GGACAATATT CAGCAACACT TCCTCCTCTC TGACCGGTTG GCCAGGGACT ATGCAGCCAT 600
55 CGTCTCTTTT GCTAACAACC GCTTTGAGAC AGGGAAGAAA AACTGCAGT ATCTGAGCTT 660
CGGTGACTTT GCCTTCTGCG CTGAGCTCAT GATCCAAAAC TGGACCCTTG GACCCGTCGA 720
CTCACAGATG GATGACATGG ACATGGACTT AGACAGGAAT TTCTCCAGGA CTTGAAGGAG 780
60

CTCAAGGTGC TAGTGGCTGA CAAGGACCTT CTGGACCTGC ACAAGAGCCT GGTGTGCACT 840
GCTCTCCGGG AAAGCTGGGC GTCTTCTCTG AGATGGAAGC CAACTTCAAG AACCTGTCCC 900
5 GGGGGCTGGT GAACGTGCCG CCAAGCTGAC CCACAATAAA GATGTCAGAG ACCTGTTTGT 960
GGACCTCGTG GAGAAGTTTG TGGAAACCTG CCGCTCCGAC CACTGGCCAC TCAGCGACGT 1020
GCGGTTCTTC CTGAATCAGT ATTCAGCGTC TGTCCAATCC CTCGATGGCT TCCGACACCA 1080
10 GGCCCTCTGG GACCGCTACA TGGGCACCCT CCGCGGCTGC CTCCTGCGCC TGTATCATGA 1140
CTGAGGTGCC TCCCAACGTC CGCCACGCT GACAATAAAG TTGCTCTGAG TTTGGAGACT 1200
15 GGTCTCGCT CCGGGGAGCA AGTGGGGGCG GTGCAGATGT GCCTGTGTCT GTCTCTGAGC 1260
ACCTGGTGTG CGTGATACAAG GATGGATGTG TNCNGTGGCT CCTTGGGAAC TGAGACATAT 1320
CTCAGGAAT GGTGTCTGTG CTCAGCCCAT CCACCAGAAG AGTCTGCTCA CAAAAAAAAA 1380
20 AAAAAAAAAA A 1391

25

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1569 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

35

GGCAGGAGTG GAGATGGCTG CGGCCGTGGC GGGGATGCTG CGAGGGGGTC TCCTGCCCCA 60
GGCGGGCCGG CTGCCTACCC TCCAGACTGT CCGCTATGGC TCCAAGGCTG TTACCGGCCA 120
40 CCGTCGTGTG ATGCACTTTC AGCGGCAGAA GCTGATGGCT GTGACTGAAT ATATCCCCC 180
GAAACCAGCC ATCCACCCAT CATGCCTGCC ATCTCCTCCC AGCCCCCACC AGGAGGAGAT 240
AGGCCTCATC AGGCTTCTCC GCCGGGAGAT AGCAGCAGTT TTCCAGGACA ACCGAATGAT 300
45 AGCCGTCTGC CAGAAATGTG CTCTGAGTGC AGAGGACAAG CTTCTTATTG CGACACCAGC 360
TGCGGAAACA CAAGATCCTG ATGAAGGTCT TCCCAACCA GGTCTGAAA GCCCTTCCTG 420
50 GAGGATTCCA AGTACCAAAA TCTGCTGCCC CTTTTTGTGG GGCACAACAT GCTGCTGGTC 480
AGTGAAGAGC CCAAGGTCAA GGAGATGGTA CGGATCTTAA GGGACTGTGC CATTCCTGCC 540
GCTGCTAGGT GGCTGCATTG ATGACACCAT CCTCAGCAGG CAGGGCTTTA TCAACTACTC 600
55 CAAGCTCCCC AGCCTGCCCC TGGTGCAGGG GGAGCTTGTA GGAGGCCTCA CCTGCCTCAC 660
AGCCCAGACC CACTCCCTGC TCCAGCACCA GCCCCTCCAG CTGACCACCC TGTGAGACCA 720
60 GTACATCAGA GAGCAACGCG AGRAAGGATT CTGTCATGTC GGCCAATGGG AAGCCAGATC 780

	CTGACACTGT	TCCGGACTCG	TAGCCAGCCT	GTTTAGCCAG	CCCTGCGCAT	AAATACACTC	840
5	TGCGTTATTG	GCTGTGCTCT	CCTCAATGGG	ACATGTGGAA	GAACCTGGGG	TCGGGGAGTG	900
	TGTTTGTAC	TTGGTTTCA	CTAGTAATGA	TATTGTCAGG	TATAGGGCCA	CTTGAGATG	960
	CAGAGGATTC	CATTTTCAGAT	GTCAGTCACC	GGCTTCGTCC	TTAGTTTTC	CAACTTGGGA	1020
10	CGTGATAGGA	GCAAAGTCTC	TCCATTCTCC	AGGTCCAAGG	CAGAGATCCT	GAAAAGATAG	1080
	GGCTATTGTC	CCCTGCTCC	TTGGTCACTG	CCTCTTGCTG	CACGGGCTCC	TGAGCCCACC	1140
15	CCCTTGGGGC	ACAACCTGCC	ACTGCCACAG	TAGCTCAACC	AAGCAGTTGT	GCTGAGAATG	1200
	GCACCTGGTG	AGAGCCTGCT	GTGTGCCAGG	CTTTGTGCTG	AGTGCTGTTA	CATGTATTAG	1260
	TTCTTTACT	GCTGACCACA	TTGTACCCAT	TTACAGAGA	AGGAGCAGAG	AAATTAAGTG	1320
20	GCTTGCTCAA	GGTCATGCAG	TTAGTAAGTG	GCAGAACAGG	GACTTGAACC	AAGCCCTCTG	1380
	CTCTGAAGAC	CGCGTCTGA	ATTTCTTCAC	TAGAGCTTCC	TCATCAGGTT	ACCCAGAAGT	1440
25	GGGTCCCATC	CACCATCCAG	GTGTGCTTGG	ATGTTAGTTC	TCCACCCTCG	AGGTGTACGC	1500
	TGTGAAAAGT	TTGGGAGCAC	TGCTTTATAA	TAAAATGAAA	TATATTCTAA	AAAAAAAAAA	1560
	AAAAAAAAA						1569
30							

(2) INFORMATION FOR SEQ ID NO: 46:

- 35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1924 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

	GGGCCCCCCC WCGWKTTTTT	TTTTTTTTTT TTTAATTAGG	ATAATGCCTT	TATTAACGAG	60
45	AATGAAACGT TCATTCCTCC	TTCCACTCCT TCTCGTTGGT	TTTCTGGACA	CAGCTCACCT	120
	GATCCTGCTA GAAACGTTGT	CAGTCTGCTT GTGGCTTCCC	TCCTTGATTG	ACTCACGCTG	180
50	TGTGATGTCT TGAGAAGTAT	CTATCCACTT CATGTGAATG	AGCACTCCAA	TATCAGCCAA	240
	CATCAATCAT TCTTACCTAA	AGAATAATAA GAAAAAGTTA	ATATAAAAGA	CAAGGGTATA	300
	AAATAAAGGT TTGAAAATGC	TAGTCAACTT CAAAATTTAA	AGAGTAAAAA	TCCAGAGATA	360
55	AAGATTGGGG GTAAGTTACA	GCATAAAAAA ATAGGAAGAA	ACTTCATGGT	GGGGGGGAAA	420
	TCTAAAATTA TTCTTACATA	AAATAAGTAG ACACCTGAAT	TAGAATGAAA	ACTGTATTTT	480
60	CTTTAAAATG TAAAGCCTG	ACTCTCAGTT TCACCAGTCT	GAGCACAAGT	TTGACTGCAA	540

CCCAAAATAT ACTATCCCTT ATGTGAAGGT ATGTGACAAC GTTGACCTCA CCAAATGAGT 600
TTTAACATCA GCTCTTTTTT CATATGAAAG CACATACCCT GCTCCCCATT CAAGTATGTC 660
5 TTCCATTGTC AGGCAGGCTG ACCACCTTCA GCAGGAGTCC TCCAAGAGTG CCCAACTCCC 720
CTTCCACAG TACACAACGC TGTAGTTGTT GTCCTGCAAT CCTTTGTATT TACCTCATTC 780
10 TTTCCCATCT AAGTCTTCAC TGAGTTTAA AGTTAGGGCT GGAAAAGCTA TGCCTTACTG 840
GGACAGCAAG GAACCAATTT TTTCTGAGG GAGAAGACAT TCACCTTCAC TATATGCCTG 900
GCAGGGCCAC AGTGCACAAA ACAAGATCA GCTTCATTC AAGTCCAGG TTTTCTTCC 960
15 TCCCTGAATG ATTACTGCAA AGGTATATG AAGTAAGAGT TCCCTGTTGC ACATGTACCA 1020
TCCATAAGGG ATACTATATC GTTTTGCAAT CTTCCTCCCA TTCTCCACAT TGTCTATCT 1080
TAAGTCCAAG CCCTTTTCAC TCTCAAAAAA AAAAAAAAAA TATTTTTTTC AGCACTGGTG 1140
20 TTCAAAAGCA ACGTTTTTAT GGTAAATGGT TTACCAGCAA CTGTTGAGAT TTCCAGTTGA 1200
GTCTTAAAAA TTGCCAATCA TTATCTAGCA GCAATGACAG ATGATTAGGA GCAGTCAAAT 1260
25 CCTCTGAATT CTTCCTTAA TAGGCAGCCA TTTGAGAACT GCACTAGCTG ACATCACTAA 1320
AACATTATCA GCTAAAGCCA AAACCAAATA AAGGCCCAGA CCAACATCCT GGCTCTCTAA 1380
AACCTGTCCA AAATCATTA GTGAAAGCA GTAAATGCAG GACTGTGGAT CATGTCCTG 1440
30 CAGCTGACAA TGATTAACAA TAGGAGACAT GCAACCCCA TTAAGGTTAA AAGTCCAAAA 1500
CTAGTCACAC GCATCTCTTT ATTGGGGAAA AGTGAGACTA TTATGCATTC TTGGTAGGTT 1560
35 TGCAACCTTG CATGAAGAGC ACCCATGCA TTTCTTTCAT CTTCAGAAA GCACCGGTAT 1620
CTGTTCCAAG GGCTAACAG TACGAAAATA CATCTGGCA TCACACCTCT GAACCCAAGA 1680
CTGTCTCAT TAAAAATAAT TTTGGTTTGT AACAAATTA TGAAATACAA TGCAAGCACC 1740
40 TCGGTATAGC ATTATTACTG AAACCACTTA ATTCCAGCT TTTTGAGTTT TTTAAAAAAA 1800
CCCACTGCAC TAAGATTCAC AATTCATTGC TACATACAAA TTAAGCTAG TAAGAACACA 1860
45 CTAACGTCAC AAGTTTCTCA TTCTAAAGTG CAAAAGCCTA ATCATCTGAA AGTGAACAGG 1920
GTAA 1924

50

(2) INFORMATION FOR SEQ ID NO: 47:

55 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 475 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

5 TGGTGTGGG CCCAGAAAMC AAGGGACCAG TGAAAACAMC CCCAGAGACT TGTATCCGCC 60
AGGAAAGCCA TTGCCAMTYC TGAGCCCTTG AAGGGCAAGG AGGGAAACAG TGTACCAGA 120
GCCCAGTAAG AACTGCTGTC ATGAAGGAGG GGCCACCTTG TAAGAGACAT CATTACTACC 180
AGAACTGTGG TGCCAAATTG CTGGTGTCTC TCTTTGGAGA AACCAACCAG ATACATCTGC 240
10 TGGAGACCCA GGTGGGCACA GAGAAGGGTG GAGAGAGAAT CTGGGAAGAG AAATGGAGAA 300
TAAGCAGCAC AGTGTTATTC ATTTCTGTAA ATTCTTATGT AGAAGGCTCA GTGTTAGAAA 360
TAAAGTTATT CTACTAGTTG CAAGTTAAGT GTTCTGTGTT GTTCTGCTTT CCTGTTAGCA 420
15 TAAGTAAACT CCTTTGGAA CTACACAGGT ATGTCTCTCC TTCAACATGT GTGAA 475

20

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:
25 (A) LENGTH: 346 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

AAGGGACAGA GACCTGGATT CAGATCTCAT TTTACAATGA AGACCCCAAT GCAGAAAGTC 60
ATGTCTGAAA TTCTGAGCTT ACTCTTCTGC CTGCTGGGAC CTGCTCTGGA TGAGAGAAGG 120
35 GAGGAAAAGG ACTAATCAGA GGAGCCAATG AAGTCACTCC ATGAGTTTCC TGAACCTGTC 180
CCAGCTAGAG ATTAACGTYT GACCWTCAAC GTAGGACACT GTGCAGATGG CTACTTGCTG 240
GCCACATGA AGACCAAAGC CAGGACCAAG CCCCMASCCT GCTWAACACG GCAGARTCTT 300
40 GCCCAGCCMA CYTCTGTGAR AATCTGCTTC CCTCCACAGC TGACCC 346

45

(2) INFORMATION FOR SEQ ID NO: 49:

(i) SEQUENCE CHARACTERISTICS:
50 (A) LENGTH: 1366 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

TAGGTGTCAG CCGCCACCCC CCCCCATAT GCAGATTTAC TSGGCATGGT AGTGGCCAGC 60
TTCTAACACA GCTGGTATTT CAAGTCTCCT GGGACCTCAC TCAGGAATGA TACCCCTCA 120
60 GTAGAAGCAG CAGGTGATCT TAACTCCTTT CAAAGAGCAG GCCTGTCTGG GAAGCCATGT 180

5
10
15
20
25
30
35
40

CCTCAGCAGG CACAGCAACC CCTCTGGAAA TGGATCACAA ACTCACTTCT CAGCCAGGCA 240
GGCCAAGCTT CTATTGTAAC AGTAGGCACA GTATAGTCGG ATCATCACAT CAGCTGGGTT 300
TTTGTTTAG TCATCTAGAG TCGTCTGGAC TAAAGGTCTT TCAGGTCTCC TTGCCCTGTG 360
AGTGCCTGAA CCTCCCCACC CGAATTCCTT CAGTTGTCTT GAGCCTCATG TCTCTCCTGG 420
TGGTGGGCCA GGCCCTGCA TGGGAAGGGA GCCTGCTGCG GGGCAGGCCA GCTGGGGGTG 480
CTCACCTATG CGCAATGANA GTTATGAAG GACTGGTTGT TGATGTTGGT GAGCGTATCC 540
TTCATGGCCA GCGCGAAGTC GGCCAGGTCA GCCAGGTGCT GCCAGCGCTC TCTCTCGGAC 600
TTGTCTTCTT GTGCCAGGGG ACCGTGGAGA AAGTGTCAAG GGCCGCTCAC TGCAGCAGCC 660
TGCTCTGCTG CCTTCCCTGG CAGTGTCTG GGGGTGGATT CCCTACAMCT AGATGTTCAA 720
GGCCTTACTT TTCCTCCAC AAAGGAGTCG CAGCCACGCT AGCTCTGACT TGCCACTGTG 780
ACAAAGTTCA CGTAGCAGGT CTAGGCAAAG ACTGGGCAAT TGAGCAGAGG AGACGGACCT 840
GTGAGTCTGA CCRYGAGSCG GRCCCTTCA CCTTGGCTGG GCTGGTCTG GTCCTTAGGT 900
TTTGTGAGT TGTCCCTGTT TGATCCCTC AACTAGGTGA TAAGCACTGG AGGGGGATGA 960
CCCGCCTTGG ACGTGTCTT TTAACCTCAT CCATATAATA GGGCCGTGGG ATGGTTGTAG 1020
AGGTAAAGCA GGATGATGGT GTTTAAGAC CAGAGCTTGG GACCAGGGCT CCTACACCTA 1080
ATTTTCTCTC CTGGTAGCTG AACAAAGTC TAAATTAGCT TAACAAAAGA ACAGGCTGCC 1140
GTCAGCCAGA GTTCTGAAGG CCATGCTTTC AGTTTCCCTT GTTGACAATT GCTCTCCAGT 1200
TCCTATGAAA GCACAGAGCC TAGGGGGCC TGGCCACAGA ACACAACCAT CTAGGCCTG 1260
AGCTGTGAAC AGCAGGGGGT TGTGTGCTG TTCTGTTTCT CTGCTTGCCG AACTTTCTCA 1320
ATAAACCTTA TTTCTTATT ATAAAAAAA AAAAAAAA AAAAAA 1366

45 (2) INFORMATION FOR SEQ ID NO: 50:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 1405 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

55 GCAGTAATTC CTGTTAGCCA CTGCATCCAC CAAACTAGT TTATTTTTC CCTCAAATTC 60
ATGATTTTTC CGTCTGTAC AAAGGGAAT TTGCTGATAG CTCTTTGGT CCCACTGTTC 120
CATTTTATGC TAATAGATTC CATTCTAGG CCCAGCCGTC TCTTGACTGA TGGTGTTCCT 180

	TTTAACCCCTT GGCATGTATA ATAGAATTTT GGTGAATGAA AGAACCCAAA TAGGCCAGAT	240
	AGTCCCCCCA GGCCTGATA TCCATAAAG GCTTGGGAAT GCATTATGTA ATTGTCCTTA	300
5	GTCTTTTGTG TGTTTTAGAA AAAAAAACA AGATGGGCTC AGATGGATGC CTACGTAAAA	360
	ATGGTTCCTA GCTGTGTA CTAACTTTT CTTTGAATG AGTAGTGAAA GGAAGGAGGA	420
	GGAAAGGAAA TTAAATGTCC TTCTAGTATT CTCTGGACTC AAGTCTGACA TATGAGATAA	480
10	TAACCTATAT TGAAATGCCA AGAATTGTAT CTGAAACAAG AGAACAGTTT GACACATTTA	540
	TCATGCCTTC ATATTACATA TTAAGTGAAA CCAATTAATA AACATATGAA ATATCCATTG	600
15	CACAAGGCAA AGGCACCTAA ACCTTTTGTT TCTTTTCTA CATAGCAGAA ATTGATTTTT	660
	TTTTTATTTT TTTAGGGGAA CCTATATAAT TATGACCCAG TGATGTCTTT TGGTGACTTA	720
	AGCTTATGAA TTCAGGTAC AATGAGTTG ATTCTAGATG GTTACTACCT TGAAAAGGAT	780
20	GTGGTGCTT TATGTGACAC GAGCCAGAGC CTGCTGGGGA ATAAACAAAG CAGGTTCAT	840
	GCCAACACCA ACTCGTAGCT TTAGTGGGCA GATGGGGAGT GGTTACAGA CTTCCCAAAA	900
25	TGTGGGGGCT TTGGGATTTT CCACACCATC CCACGTGTGT TGTTCAATTCT TCCTCTTTTC	960
	ACACTCTTGG ATGGATWATT TGRAAATGGT GRAAWYMCY YYKRAATTTG CCCAATAGCC	1020
	WTGRGCCACC ATTCTTWATG ACACCATAAC CAAATAGTTC CWTAAATGTTG AAATATTAGA	1080
30	AACCTGTTAC CAGCCYKMA KIWACCCWA WTTTCCCAT GTTTGTGGAA TTGATATTGA	1140
	AATAGCAGGG CTAAGGAATT ACTGGCAAGT TTTAGCCTGT GGGTAATACC TTAGGGTTAT	1200
35	TTAAATATTT GTAATTTTAT TTAAATGTTT ATGAATGTTT GAAAGGAACA AAATTATCAG	1260
	GGATGGCTCT TTGCCATGGG TCTTATTTTC ACCCTCTTTT CTGTAAGAAA AAAGAACAAT	1320
	GTCTTAATGT ATTTTAAAG TTTTGTGTAT AGTTTCTAAT TCCAATTTTA ATAAAAGTTT	1380
40	TWRTAAAAA AAAAAAAAAA AAAAA	1405

45

(2) INFORMATION FOR SEQ ID NO: 51:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 504 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

50

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

55

CGGATTTTCT AGGACCCCAA AAAAAAAAAA AGGGNAAAAA AAACCCNCAA AACCANCCAA	60
AACCCCAAAA AAAAAAAAAA TCCACAAAAA CAAAAAACT ATAAAAAGA AAGAATTAAA	120
AACTTTCAGA GAATTACTAT TTACTTTATT AACTTACGGA TTTATTATAT AAATATATAT	180

60

TCACCTAGCA ACATATCTCT GCCGTCTCTC CTGCTCTCAT AATGAAGACA TAGCCGATTC 240
TCTGCCCCGG CCCCTTGCTG ATGCTCCTCC GGGTCTGCGT CGGGCGTGGG TCTCTGGGA 300
5 CCCTCCAGAG GTGGAGGTGG GCTGATGGCC TGGCTGCCCTG GTGGTTGATG GTTTTGCTCC 360
CCCTACCTTT TTTTTTTGAG TTTATTCTGA TTGATTTTTT TTCTTGTTT CTGGATAAAC 420
10 CACCTCTGG GGACAGGATA ATAAAACATG TAATATTTTT AAGAAGGAAA AAAAAAAAAA 480
AAAAAACTNG GGGGGGCCCC CGAA 504

15

(2) INFORMATION FOR SEQ ID NO: 52:

(i) SEQUENCE CHARACTERISTICS:
20 (A) LENGTH: 777 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

NAAGTATCTT GGCCAGTTTA TTACAGAGGA CGATAAATGA TTCCATGTGG ATAGGCATA 60
ACATACAGAG AATGAGACTA TGCCAGAAAT GGGAGGAGGC ATTTGAAACA ACATGAGTAT 120
30 CTCAGGACA GATGGATTGA TTCTGCTATT GGTAGGCCTG GAAGCAANGG TCAGAAGTAG 180
CAAAAAATGG ATACCAAAAG CACTATTWGT CACCCAAGCT AAGTGAATA GCTGGCCCAG 240
35 TAGGAGAAAT GCAGTTTTTG CTCTACACTA AGTTCTCCAA CTCTTGATAA GCCTCCAAAA 300
ACAAATGTTA GGGGAAAAAA ACGCAGCTGG TTATGAAAAG ATATATCTCA TTTCAATAAA 360
AAATCAATGT CAATGCTGTT AATAGAATCC TTTATCTTC AGGACAGAGG CAATGCCCTA 420
40 AACAAACACC AGCTCAAGAG CCTCTGATGC CAACCTAGAG GGTACCCAAA CACAACTTA 480
GCATAGAGGT AAGAATCTCT ATGTCTTTTG GTGAGGCAA AGCCATTGG TTGGTACTTC 540
45 ACAGGAACAT CTTTCTACCA AGTCTTCATC ATATGGTATG TGCCACGAGT CTCCAGTTGT 600
TTGACCACT GTGTCATAGC TGAGAATACG CTGAAAGGTT AGTTTGTATC CTGGAAACCT 660
ATTTACAATT GCCAGTGAT GTCCCTGCTG CCACTTAAAA AAGGCTTGGG TCTGGCATAG 720
50 GCAGAMAGGC CTGTGGTCCC CTCGTGCCGA TTCTNGGCTC GAGGCCAATT NCCTTAT 777

55

(2) INFORMATION FOR SEQ ID NO: 53:

(i) SEQUENCE CHARACTERISTICS:
60 (A) LENGTH: 602 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

5
ATGACTACAG TGTATATACC TCCAATCTTT GCAGGTGGGC ATGGAACACT GCTTGATATCA 60
CTCTGTGCAC GGTATAAATC CATATATCCA CAAAACACA CATCCATCCA TCAACATATA 120
10 CATGGTTTGG GATGAGCAGG TCAATAGTTT TGAGAGGGAG TTTGTTCCCTT TTTTITTTCT 180
CATTATACTC TTAAATTGTT GTCAGTTATC AAACAAACAA ACAGAAAAAT TGTITGGAAA 240
AACCTTGCAT ACGCCTTTTC TATCAAGTGC TTAAAAATAT AGACTAAATA CACACATCCT 300
15 GCCAGTTTIT TCTTACAGTG ACAGTATCCT TACCTGCCAT TTAATATTAG CCTCGTATTT 360
TTCTCACGTA TATTACCTG TGACTTGTAT TTGTTATTTA AACAGGAAAA AAAACATTCA 420
20 AAAAAAGAAA AATTAAGTGT AGCGCTTCAT TATACTATTA TATTATTATT ATTATTGTGA 480
CATTTTGGAA TACTGTGGAA GTTTTATCTC TTGCATATAC TTTATACGGA AGTATTACGC 540
CTTAAAAATA CGAAAATAAA TTTTACAAGG TTCCGGTTT TGTGGTGGAA AGAGTAAATT 600
25 GA 602

30

(2) INFORMATION FOR SEQ ID NO: 54:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 1749 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

40 AGTCACTGAC TTGGAGCCGC TCGGGGAAG TCCCGCCCAG ACAGGCGGTG GTGGGAATG 60
CCTCACTTCA GTTTGAAGAG GGTCCGATC CAAAGGGTT AAAACGAGCG AACCCCGATC 120
45 CCCGACCACA CTTCCCGCCT CCCTAAAACG CACACCCCGC TAGCCATGGG CAGCCGCGAC 180
CACCTGTTCA AAGTGCTGGT GGTGGGGGAC GCCGAGTGG GCAAGACGTC GCTGGTGCAG 240
GATTATTCCC AGGACAGCTT CAGCAAACAC TACAAGTCCA CGGTGGGAGT GGATTTTGCT 300
50 CTGAAGGTTT TCCAGTGGTC TGAATACGAG ATAGTGCGGC TTCAGCTGTG GGATATTGCA 360
GGGAGGAGC GCTTACCTC TATGACACGA TTGTATTATC GGGATGCCTC TGCTGTGTT 420
55 ATTATGTTTG ACGTTACCAA TGCCACTACC TTCAGCAACA GCCAGAGGTG GAAACAGGAC 480
CTAGACAGCA AGCTCAGCT ACCCAATGGA GAGCCGGTGC CCTGCCCTGCT CTGGCCAAC 540
AAGTGTGATC TGTCCCTTG GGCAGTGAGC CGGACCAGA TTGACCGGTT CAGTAAAGAG 600
60

AACGGTTTCA CAGGTTGGAC AGAAACATCA GTCAAGGAGA ACAAAAATAT TAATGAGGCT 660
ATGAGAGTCC TCATTGAAAA GATGATGAGA AATTCCACAG AAGATATCAT GTCTTTGTCC 720
5 ACCCAAGGGG ACTACATCAA TCTACAAACC AAGTCCCTCA GCTGGTCCCTG CTGCTAGTAG 780
TGTTTGGCTT ATTTTCCATC CCAGTTCTGG GAGGTCTTTT AAGTCTCTTC CCTTTGGTTG 840
10 CCCACCTGAC CATTTTATTA AGTACATTTG AATTGTCTCC TGACTACTGT CCAGTAAGGA 900
GGGCCCATTG TCACTTAGAA AAGACACCTG GAACCCATGT GCATTTCTGC ATCTCCTGGA 960
TTAGCCTTTC ACATGTGCT GRCTACATT AGTGCCAGTT AGTGCCTTCG GTGTAAGATC 1020
15 TTCTCATCAG CCTCAATTT GTGATCCGA ATTTTGTGAG AAGATTAGA AATCAGCACC 1080
TGCGTTTTCAG AGATCATAAT TCTCACCTAC TTCTGAGCTT ATTTTTCAT TTGATATTCA 1140
20 TTGATATCAT GACTTCCAAT TGAGAGGAAA ATGAGATCAA ATGTCAATTC CCAAATTTCT 1200
TGTAGGCCGT TGTTCAGAT TCTTTCTGTC TTGGAATGTA AACATCTGAT TCTGGAATGC 1260
AGAAGGAGGG GTCTGGGCAT CTGTGGATTT TTGGCTACTA GAAGTGTCCT AGAAGTCACT 1320
25 GTATTTTGA AACTTCTAAC GTCATAATTA AGTTTCTCTT GTCTTGGCAT CAAGAATAGT 1380
CAAGTTTTC GTCCGGGCAT GGTGGCTCAT GCCKGTAATC CCAGCACTTG GGGAGGCCAA 1440
30 GGCAGGCGGA TCACATGAGG CCAGGAATTC GAGACCAACC TGGTCAGCAT GGCAAAACCC 1500
CGTCTCTACT AAAAGTACAA AAATTAGCCA GCGTGATGG CACGTGTCTG TAATCCAGC 1560
TACTCTGGAG ACTGAGGTGG GAGAATCGCT TGAGACTGGG AGGCAGAGGT TGCACTGAAC 1620
35 CGAGATCATG CCACCGCACT TCAGCCTGGG TGACAGAGAA GGACTCCGTC TCAAAAAAAA 1680
AAAAAAGAG AAAACTCGAG GGGGGGCCCG GTACCCAAAT CGCCSTGATA GTGATCGTAW 1740
40 ACAATCNA 1749

(2) INFORMATION FOR SEQ ID NO: 55:

45

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1896 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

50

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

55 AAAGAGATGG GCTCTTTATT TTCTCGAAAA ACCAATTGG AGTTACTCAT TTTTCATAA 60
CATTAATTT CTACAGTGA ACTACATATT GTCCATAAGT GCTTCATCAG GACTCATCGC 120
CCTCCTGTCT ACTGGCTCCA AATAGACCAT GTCAGCTTCA CCCCTGGCT TTGTGTCTAT 180
60 GGGTGGCCTG TGGTATATGG AAAAGTAGCA GGGTGGTCAG GGTGGGAGAC ACAAGATGTT 240

	TTTATAGTCT AGAGCCTTTA AAAAACCAG CAGAATGTAA TTCAGTATTT GTTTATTGGC	300
	TGTTTTTTGA CAGATTGTTG AAATTAAATG AATTGAAAGG GAAACTCAGA GTAGTAGGAC	360
5	GTTTATTAAA AGGAAAAAAA TGTCTTGCAA TGTGCTGTAA TCACAAGAGG AGAAAAAATAC	420
	TTGTTTCCTT GATCTGTCAG AGGTCACAGT AACCTGGGCC GAGCTGTTAT TATTTATTAT	480
10	ATAATAGTAG TAGGAAGTTA ATAACCTGGT CTCTGTGTTT CAAGCACAAT ATTACAACCT	540
	CTTTTGAACC GTAAATATCA GAATGAATCC TCTTCCAGG GGATTGAACA GAAGCTTAAT	600
	GTTTACAAGT GTTTGAATTT GTGATCTGAA ATAACACAAA ATTAAAAACA TGATTTCTCT	660
15	AATTTTCCAA CTAGAGGAAG AGAACTTGT GGAAAAGTTC TTTTTTTTC TTTTTTTTTT	720
	CTTAAAGAAG GGCAGCCAAG GTAGTAACCT AAAAATAGTG CCCAGGCATA TGAGAGTTGT	780
20	CCTACGAGGT TAAAGAACAC ACTGTTCCAC TGTATGGCTT TGGCCCTGAG TGGCCAGGGA	840
	GGTCAACTTG ACCCTGCCAT GTTGGTTTGA CTTACTAAGA CACAGGAATC ATTGTTTTC	900
	TTGACCAGGG TCTCACACC TGGAGGAATG TTAAGTAAGA GAAAGAACCT CTTTCTGAA	960
25	TATTGACATG TAAAAGACCA AAGTAATTTT TCTGAACCTC TGCAATCTCG AGAACTCTCC	1020
	AAGGAATTTA CAGTGATTTT AGTGCTTGTC AGCATTTTTC CATGAGGACT TTCATACATT	1080
30	TGACTCTTTA GTTCACAGGT TCCCATTGAT TGTGAGCAAG ATATTTATCT CTTTAGCCCT	1140
	TGGGGATCCA GCTGAGAGCA ATCTCTTGCA TTTTMTTACC CGTGATGTA CAGATATCAT	1200
	TTCTTGTTGA TGCCATGACT TGAAAAAGTT TGGGAAGCTC TTTAGCAATA TCAGCTAAAA	1260
35	GGATATGAAA TCACAGGTGA TAGCAGTTGT CATTCAGTAA TTTCTTACAA GCAGACCCCC	1320
	AAAGGAAATA TAGTCCTAAT CTTTACTATC CACTTCTAAA TTTAATGTGA ATTTCATACA	1380
40	TGTTATTAGT TGTTTTCTTT ATAATTTTAT AAAAATTATT CATCGGGAGT TTAACTTCCA	1440
	CTTCCATGCT ATCGGATGTG TTGGGCTCCA TGCAAGAACT TGAAGAAAA ACAGGCAGGA	1500
	ATGCATTTGC ATAATGACCC AGATCATCAT TTTCTGCAAC TGAGAATTAT ATTTTCATCAT	1560
45	TGCTTCTAGA AGTCTGCAAT TCTTTACTTT TCTTTGGTGC ATTATPATCT AGGTGCCATC	1620
	ACTGGATAAT GTGGAGTGAC TAGAGAAGTC AYATATCACT GTAAGGTACA GTTAGGGGTA	1680
50	ACACITTAGA GGTTTATTAT TTTTAAAAA CTTTCTTGA ACTCCTGGGC CAACATGGGT	1740
	GAAACCCCGT CTCTTACTTT AAAAATACCC AAAATTAGGC CAGGGGCGTG GATGGGTGGG	1800
	GTGCTGTGTA ATCTTCAGCT ACTTNGGGA GGGCTTGAAG CCAGGGAGGA ACTGCCCTGG	1860
55	ANCCCCGGG NGGGCCAGNA GGTTCGCCAG TTGAGT	1896

(2) INFORMATION FOR SEQ ID NO: 56:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 1753 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

10 TCTTTTAA ATAGACATTT GTGGGCTCA CACAATATAT GAAATAGTAC CCTCTAAAA 60
AGAGAAAAA AAAATCAGGC GGTCAAACCT AGAGCAACAT TGTCTTATTA AAGCATAGTT 120
15 TATTTCTACTA GAAAAAATTT AATATCAAGG ACTATTACAT ACTTCATTAC TAGGAAGTTC 180
TTTTTAAAT GACACTTAAA ACAATCACTG AAAACTTGAT CCACATCACA CCCTGTTTAT 240
TTTCCTTAAA CATCTTGAA GCCTAAGCTT CTGAGAATCA TGTGGCAAGT GTGATGGGCA 300
20 GTAAATACC AGAGAAGATG TTTAGTAGCA ATTAAAGGCT GTTGCACCT TTAAGGACCA 360
GCTGGGCTGT AGTGATTCTT GGGGCCAGAG TGGCATTATG TTTTACAAA ATAATGACAT 420
25 ATGTCACATG TTTGCATGTT TGTTCGCTTG TTGAATTTT GAACAGCCAG TTGACCAATC 480
ATAGAAAGTA TTACTTTCTT TCATATGGTT TTTGGTTCAC TGGCTAAGA GGTTCCTCAG 540
AATATCTATG GCCACAGCAG CATACCAGTT TCCATCCTAA TAGGAATGAA ATTAATTTTG 600
30 TATCTACTGA TAACAGAATC TGGGTCACAT GAAAAAAT CATTTTATCC GTCTTTTAAG 660
TATATGTTTA AAATAATAAT TTATGTGTCT GCATATTGCA GAACAGCTCT GAGAGCAACA 720
35 GTTTCCTT AACTCTTTCT GACCAATAGT GCTGGCACC GTCCTTCCTC TTTGGGAAGA 780
GGAAAGGCTG TGTGAACATG GCTAACAATC TTCAAATACC CAAATGTGA TAGCATAAAT 840
AAAGTATTTA TTTTATGCCT CAGTATATTA TTATTTAATT TTTAGGTAA TGCCTATCTC 900
40 TTGGTCTATT AAGGAAAGAA GCAATCAGTA GAGAAATCAG GATAGTTTGT TTTAAATCTT 960
TGCAGATTAC ATGTTTATAC AGTGGCCTGC TATTGAGGAA AGGTATTCTT CYATACAACT 1020
45 TGTTTAACC TTTGAGAACA TTGACAGAAA TTATGCAATG GTTGTGTGAG ATACGGACTT 1080
GATGGTGCTG TTTAATCAGT TTGCTTCCAA AGTGGCTTAC TCAAGAGGCC CTAAGACTGG 1140
TAGAAATTA AAGGATTTCA AAAACTTTCT ATTCTTTCT TAAACCTACC AGCAAACCTAG 1200
50 GATTGTGATA GCAATGAATG GTATGATGAA GAAAGTTTGA CCAAATTTGT TTTTGTGTG 1260
TTGTGTGTGT TTTGAATTTG AAATCATTCT TATTCCTTTT AAGAAATGTTT ATGTATGAGT 1320
55 GTGAAGATGC TAGCGAACCT ATGCTCAGAT ATTATCGTA AGTCTCCCTT CACCTGTTAC 1380
AGAGTTTCAG ATCGGTCCT GATAGTATGT ATTTCTTTAG TAAGAATGTG TTAATAATTAC 1440
AATGATCTTT TAAAAAGATG ATGCAGTTCT GTATTATTG TGCTGTGTCT GGTCTAAGT 1500
60

GGAGCCAATT AAACAAGTTT CATATGTATT TTTCCAGTGT TGAATCTCAC ACACTGTACT 1560
TTGAAAATTT CCTTCCATCC TGAATAACGA ATAGAAGAGG CCATATATAT TGCCTCCTTA 1620
5 TCCTTGAGAT TTCACTACCT TTATGTTAAA AGTTGTGTAT AATTGTTAAA ATCTGTGAAA 1680
GAATAAAAAG TGGATTTAAA TTAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA 1740
AAAAAAAAGG GGG 1753
10

(2) INFORMATION FOR SEQ ID NO: 57:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1220 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

20

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

25 GCGGAAGTTA CTGCAGCCGC GGTGTTGTGC TGTGGGGAAG GGAGAAGGAT TTGTAAACCC 60
CGGAGCGAGG TTCTGCTTAC CCGAGGCCGC TGCTGTGCGG AGACCCCGG GTGAAGCCAC 120
CGTCATCATG TCTGACCAGG AGGCAAAACC TTCAACTGAG GACTTGGGGG ATAAGAAGGA 180
30 AGGTGAATAT ATTAAACTCA AAGTCATTGG ACAGGATAGC AGTGAGATTC ACTTCAAAGT 240
GAAATGACA ACACATCTCA AGAACTCAA AGAATCATAC TGTCAAAGAC AGGGTGTTC 300
AATGAATTCA CTCAGGTTTC TCTTTGAGGG TCAGAGAATT GCTGATAATC ATACTCCAAA 360
35 AGAACTGGGA ATGGAGGAAG AAGATGTGAT TGAAGTTTAT CAGGAACAAA CGGGGGGTCA 420
TTCAACAGTT TAGATATTCT TTTTATTTT TTTCTTTTCC CTCAATCCTT TTTTATTTT 480
40 AAAAATAGTT CTTTTGTAAT GTGGTGTTC AAACGGAATT GAAAACTGGC ACCCATCTC 540
TTTGAACAT CTGGTAATTT GAATCTAGT GCTCATTATT CATTATTGTT TGTTTTCATT 600
GTGCTGATTT TTGGTGATCA AGCCTCAGTC CCCTTCATAT TACCCTCTCC TTTTAAAAA 660
45 TTACGTGTGC ACAGAGAGGT CACCTTTTTC AGGACATTGC ATTTTCAGGC TTGTGGTGAT 720
AAATAAGATC GACCAATGCA AGTGTTCATA ATGACTTTCC AATTGGCCCT GATGTTCTAG 780
50 CATGTGATTA CTTCACTCCT GGACTGTGAC TTTCAGTGGG AGATGGAAGT TTTTCAGAGA 840
ACTGAACTGT GGAAAAATGA CCTTTCCTTA ACTTGAAGCT ACTTTTAAAA TTTGAGGGTC 900
TGGACCAAAA GAAGAGGAAT ATCAGGTTGA AGTCAAGATG ACAGATAAGG TGAGAGTAAT 960
55 GACTAACTCC AAAGATGGCT TCACTGAAGA AAAGGCATTT TAAGATTTT TAAAAATCTT 1020
GTCAGAAGAT CCCAGAAAAG TTCTAATTTT CATTAGCAAT TAATAAGCT ATACATGCAG 1080
60 AAATGAATAC AACAGAACAC TGCTCTTTT GATTTTATTT GTACTTTTTC GCCTGGGATA 1140

5 TGGGTTTTAA ATGGACATG TCTGTACCAG CTTCATTAAA ATAAACAATA TTGTAAAAA 1200

TCAWAAAAAA AAAAAAAAAA 1220

5

10 (2) INFORMATION FOR SEQ ID NO: 58:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1049 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

20 TCGCGCCTGC AGACACAGCA TCTACTCAGC GTGGGTCACC TCTGTGAACA TCACTGACTG 60
CAAGCCTCCC TCAATTTCTG GTGCAGCCCA TCAGGGACCC ACAGCGCCTG GGAGGATGGT 120
GCGGATCTTG GCCAATGGGG AAATCGTGCA GGACGACGAC CCCCAGTGTA GGACCACTAC 180
25 CCAGCCACCA AGAGGTAGCA TTCCTCGACA GAGCTTCTTC AATAGGGGCC ATGGTGCTCC 240
CCCAGGGGGT CCTGGCCCCC GCCAGCAGCA GGCAGGTGCC AGGCTGGGTG CTGCTCAGTC 300
CCCTTCAAT GACCTCAACC GGCAGCTGGT GAACATGGGC TTTCGCAGT GGCATCTCGG 360
30 CAACCATGCT GTGGAGCCGG TGACCTCCAT CCTGCTCCTC TTCTGCTCA TGATGCTTGG 420
TGTTCTGTGC CTCTCTCTGG TTGGCCTTGT CTACCTGGTG TCCCACCTGA GTCAGCGGTG 480
35 ACCTCTGAGG GCTGATAGGG GTGGGTTTGT TGAGAGGGAC TTGCTGGGCC TTGGTGTGAG 540
AGCAGGCATA TTTGGAGGGG ATCTGGTGGT GCCTTGAAGG TATGATCAGA GAGGGGACCA 600
CAGGTGTGTG TTTCCCTTTT GTGTTAAGCG TGAGGCAGAG GGAGACGTTA GTCCCAGCAT 660
40 TTCCCAAAGT GTGGGTGGGT CCGTTGGTTC CCGAGATACT TTTAGGTGGT ATGGGGCCTG 720
CATTAAGTGG CACAAAATCA GAGCAAGAAA GCGATGCCCT TCCCAATTCT CTCAATCCTT 780
45 TTATGCCGAG AAGATCTCAG CTGGATGCCA ACATGTTCCG ATGCCTGTGG AAGACATGCC 840
GACGTCTCCT CTGCCTAGGG AGCAGGACTT GGGCTTAGGG CAGGTGGAAG AAATTCCAGA 900
CTTTTTTAGC ACTGTTTTTG TTTTAATGGT ATATTTTTAT TGGCTACTTT ATGTGTTAGG 960
50 ACAAGTGTA GTGGCATTCT ATTTATTGTG ACCTTTTCAA TAAATAGATT TAAGTAAAAA 1020
AAAAAAAAA AAAACTCGAG GGGGGGCCC 1049

55

(2) INFORMATION FOR SEQ ID NO: 59:

60 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1776 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

	AAAGAGGATG TGMAGCTAGA GGTCCCCGAT GGCTGGTCGG ATGGGAAGCA CAAGGCTGAG	60
10	GGACTGGATT GTAAAGGCAC TAAGTCGTTT TCGCGTGAGA ATCAGACATG GGGGACCTCT	120
	AGCTTCACAT CCTCTTTCTT TGCAGSTCTG GACATCCTGA GCCCAAGTCC CCCCACTCA	180
	GTGCAGTGAT GAGTGCAGAA GTGAAGGTGA CAGGGCAGAA CCAGGAGCAA TTTCTGCTCC	240
15	TAGCCAAGTC GGCCAAGGGG GCAGCGCTGG CCACACTCAT CCATCAGGTG CTGGAGGCCC	300
	CTGGTGCTTA CGTGTTTGGA GAACTGCTGG ACATGCCCAA TGTAGAGAG CTGGCTGAGA	360
20	GTGACTTTGC CTCTACCTTC CGGCTGCTCA CAGTGTTCGC TTATGGGACA TACGCTGACT	420
	ACTTAGCTGA AGCCCGGAAT CTTCCTCCAC TAACAGAGGC TCAGAAGAAT AAGCTTCGAC	480
	ACCTCTCAGT TGTCAACCCTG GCTGCTAAAG TAAAGTGTAT CCCATATGCA GTGTTGCTGG	540
25	AGGCTCTTGC CCTGCGTAAT GTGCGGCAGC TGGAAGACCT TGTGATTGAG GCTGTGTATG	600
	CTGACGTGCT TCGTGGCTCC CTGGACCAGC GCAACCAGCG GCTCGAGGTT GACTACAGCA	660
30	TGGGGCGGGA CATCCAGCGC CAGGACCTCA GTGCCATTGC CCGAACCTCK AANAAAAACC	720
	ATTAAAGTTA CGACGGCAGC AGCAGCCGCA GCCACATCTC AGGACCTTGA GCAACACCTG	780
	ACTGAGCTGA GGAACCAGC TCCTGGCACC AACCAGCGCC ASCCAGCAAG AAAGCCTCAA	840
35	AGGGCAAGGG GCTCCGAGGG ANCGCCAAGA TTTGGTCCAA GTCGAATTGA AAGRACTGTC	900
	GTTTCCTCCC TGGGGATGTG GGGTCCCAGC TGCTTGCCTG CCTCTTAGGA GTCCTCAGAG	960
40	AGCCTTCTGT GCGGCTGGCC AGCTGATAAT CCTAGGTTCA TGACCCCTCA CCTCCCCTAA	1020
	CCCCAAACAT AGATCACACC TTCTCTAGGG AGGAGKCAA TGTAGGTCAT GTTTTGTGTG	1080
	GTACTTTCTG TTTTPTGTGA CTTTCATGTG TCCATGTCTC CCCGCTGCCA TGCTCTCTCC	1140
45	CTTGTTCCTT TAAGAGCTCA GCATCTGTCC CTGTTTATTA CATGTTCATT AGTAGGTGGG	1200
	TAGCCCTGAT GGGGGTCGCT CTGTCTGGAG CATAACCCAC AGGCGTTTTT TCTGCCACCC	1260
50	CATCCCTGCA TGCCCTGATCC CCAGTTCCTA TACCCCTACCC CTGACCTATT GAGCAGCCTC	1320
	TGAAGAGCCA TAGGGCCCCC ACCTTTACTC ACACCCCTGAG AATTCTGGGA GCCAGTCTGC	1380
	CATGCCAGGA GTCACCTGAC ATGTTTCATCC TAGAATCCTG TCACACTACA GTCATTTCTT	1440
55	TTCTCTCTCT TGGCCCTTGG GTCTGCGGAA TGCTGCTGCT TCAACCCAG AGCCTAAGAA	1500
	TGGCAGCCGT TTCTTAACAT GTTGAGAGAT GATTCTTTCT TGGCCCTGGC CATCTCGGGA	1560
60	AGCTTGATGG CAATCCTGGA AGGGTTAAT CTCCTTTTGT GAGTTTGGTG GGAAGGGAA	1620

GGGTATATAG ATTGTATTAA AAAAAAAAAAG GTATATATGC ATATATCTAT ATATAATATG 1680
ACGCAGAAAT AATCTATGA GAAATCTATC TACAAAMWAA AAAAAAAAAA AAAAAAAAAA 1740
5 AGGAATTCGA TNTCAAGCTT ATCGATACCG TCNACC 1776

10

(2) INFORMATION FOR SEQ ID NO: 60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 443 base pairs
15 (B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

20 ACAGATAAAT AATAAATAA TAAATTAAAT TAAATAAAAA ATCTGAGCTA ATCTGAATAA 60
ATPGAGAGAT TTCACATGAA AGCCAGGATT TCTGGCTTCC CAGGAACAGT CAGAAGAGCT 120
25 AGCTAGCAAC ACTGGTCTGC TTGGCTACCT TCTTTGGAAC AACATGAAAT CTAGCTCCCT 180
TTTTTTTTTT TTTTGGCCC ACTTCATCCA TTCACATGAC CTGCCTGGCC TCTGCAGGTA 240
AGTGAGTATG CAACAAAAAT GTAGCACAGG TTTTGTCCCT GAACTACGTG GTTTCAGGTC 300
30 CAGCTCTGCC ACTTGCTAGC ATGACCTCGT GCCGAATCC NGCAGGAAGT TTTTTTTTTT 360
TTTTCAGTG CTCCAGTCCC CCTATTGGAG AATCCTGCCC CCCCTGGGA CAGAAATGTT 420
35 ACCCTGGCCC CGCGANTCCC TGA 443

40

(2) INFORMATION FOR SEQ ID NO: 61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2888 base pairs
45 (B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

50 TTAATGTTGT CAATAACCAC CAGGCCAAAC AGAATTTATA TGACCTGGAT GAAGATGATG 60
ATGGTATAGC TTCCGTTCTT ACTAACAGA TGAAGTTTGC AGCCTCAGGC GNCCTTCTCC 120
ACCACATGGC TGGGCTAAGC AGTTCCAAGC TTTCCATGTC CAAGGCCCTC CCTCTCACCA 180
55 AAGTGGTTCA GAATGATGCA TACACAGCTC CTGCTCTCCC TTCTCTATTT CGAACAAAAG 240
CCTTGACCAA CATGTCCCGG AACTGGTGA ACAAGGAAGA ACCCCCCAAA GAGCTGCCAG 300
60 CTGCTGAGCC TGTCTCAGC CCATTGGAAG GCACCAAGAT GACTGTGAAT AATCTGCACC 360

	CTCGAGTCAC TGAGGAGGAC ATTGTTGAGC TTTTCTGTGT GTGTGGGGCC CTCAAGCGAG	420
5	CTCGACTGGT CCATCCTGGG GTAGCGGAGG TGGTGTTTGT GAAAAAGGAC GATGCCATCA	480
	CCGCATATAA GAAGTACAAC AACC GG GTGC TGGACGGGCA GCCGATGAAG TGCAACCTTC	540
	ACATGAATGG GAATGTTATC ACCTCAGACC AGCCCATCCT GCTGCGGCTG AGTGACAGCC	600
10	CATCAATGAA AAAGGAGAGC GAGCTGCCTC GCAGGGTGAA CTCTGCCTCC TCCTCCAACC	660
	CCCCTGCTGA AGTGGACCCT GACACCATCC TGAAGGCACT CTTCAAGTCC TCAGGGGCTT	720
15	CTKTGACCAC GCAGCCCACA GAATTCAAAA TCAAGCTTTG AGCAGGGGAG TGAGGCAGCC	780
	AGAAGTGGGG GCAGAGGAGG GTGGCTCTGT TTCCCAAGG CAAAGCTTAT GACCAATGGG	840
	CCATCGGACT GGAGACCCTT GATTGTGGGA AGGGTTGCCA GGGATAAAGA GCTTCCTCAC	900
20	TGGATGGGAC CCGCCTTTCT GTGTGTGTGT CTGCCCTGTG CTCTTCTCTC TACGTTAACG	960
	TTTCCTGTAG TATGTTTCTT CATCTCATCG CCAAGGTAGG CTGTGTTTT TCAGTGTGTG	1020
25	CCTCCCCGAG CCTCAGCCCC AAGCTGATTT CTTATCTGGA AATGGTACAC TGAATCTCT	1080
	GGGTGGCTTT CTGTGGGCC CATGGGATGC AGCGTGGGG CTGTCTGAAG GACCCTGCTT	1140
	TTTCAGGGG CCGAGGGGCT GCCTTTCCTT TGTGTGTATT AAGCTTTTCA AACAATGGAG	1200
30	GGGATGGAGA GCCCTGGTGT CCTGACGGGA GCCAGGTCGG CCTGAGAGCT GTGCCGCTCC	1260
	TCTGTCTGTG CAGTGGAGGT GCCTGGGTGG GGAGCAGGTC TCAGGCCTCT TGTCTCTCC	1320
35	CCAGTGGCTC CAGGCCTCAC TAGTGGCAAG GGCAGGATGA GGCTGCACCG CTGGGAAGAG	1380
	TCTATCTAAG YCTTTGGCTT GGAGTCCCGT GTCGTCTCCR CCCAGAGGAA GTTCTCCAGA	1440
	GTTACCTTTT CCCTTTTCCT TGAGTTGTGC TGAATGCCCC ACCCCAGCTC TCTTTCCCTT	1500
40	CTGGGTGTCT TTGCTGGGAG GGGGCTGTGT TGTGAGCCCT CCCGGTCTC ACCTCGCCTG	1560
	GCACTTAACC ACACCCTGGT TTTGTGTAGC CGCCAGCTCT CTCTGGTTG GGCCTTTGAA	1620
45	AGGCTCAGCC TCCCATTGTG CAGTGCCTTG GTTTGGAGCT TATTTGAATG GAAGAGGTCA	1680
	GTTTGTTCCT GGCTCTCCAT TTCTGGCCTC AGTTGTCTAC AGGACAGTGG TCAGGGATGC	1740
	CTGGAGGCAT ATATCCAGCT GCCACCAAGG GGCAGTGTG GTTCCCACTT ATGTGAGTGA	1800
50	CCCCATCCAT CCATGACCAG AGGATTATTT TCCTGCCTTG GCAGAGGAGG AGGAGTCAAG	1860
	GGAGCAGGGC AGCTCTACCA GGCAAGGTGT TTCCCCAGCA TAGGCGCAGA CAGTTGGGAC	1920
55	GAAACTTCAG AGCCCAGGCA GTCCCTGAAT GACCAGGCCA GTGTTGTAC TGAGTGGTCC	1980
	CCTGCTGGTT GGGAGTGAAG AGAATCCAGG CTGGCAGAGC TGGAGCCAGT TGGGAGCAC	2040
	GGTCTGGGA GCTCTGCAA ATCAGTAGCA AGTGCTGGAA AAGGCACATG CCGAAGATAC	2100
60	TCAAGAGCTC CCAAGATTTG CTTGAGGCTA GCCCAGTGAA RAAAACCAGA GACTCATGTT	2160

5
10
15
20
25

TCCAGGGGTC AGTCTGTCAG GCAGGAAGGA CCCAGGATTT GAACCCAGCT TCAGTGTGCA 2220
GGCTCTGAGG CTGCCAGGA CGGGAAGTC CAAGGAAGGG GCCTGGTGGT GCTCCACTTG 2280
CAGTTCCTTA AAGATGCTG CTTTTATTC TCCTAACCTT TTCAAGTGGG TGCAGACTTC 2340
TCGTTAGCAG CTGGAAGACA TTCCTCCAC ACTTTTCCCT TCCTGGCCCA AGAGAGCATC 2400
CAGAAGGCAG TAGGACCTGG TTTTCAGGT ACTGGGAGCC GGGGGCTCAC TGCTTGCACT 2460
GTGCTTAGGG TAGGATGGT AAATATCCTC CCTGCATGGC TTTATCCTCC CTCTCATCCC 2520
AAAGCAGGTA TCTTCTGGTT GTCACAGAGT TTCATTGAGT CCAGCTGCAG CCACGTGGCC 2580
ATCTGGAGCT GGTGCTATAG GTGACCATCT GGTACATTGA GGGGACCTGT TTGCCTCCTC 2640
CACTCTATAA GCAGTCATCT TGGGAGACCG GGAGGAGAAG GTGGTGGGCT AGTCCTGTGT 2700
CCTCCTCCAC TTCCCATGCC TCTATGTAC CCATCTGTGT CTCTGTGCA GAAGGAGAGG 2760
AAGGGGCATT AAGAGATGAA GGGTGATTAT GTATTACTTA TCCATTTC TG AATAAACATT 2820
TGTTATTCCT AAAAAAAAAA AAAAAAACT CGAGGGGGGG CCCGGWACCC AWATCGCCSK 2880
AAAGTGAG 2888

30

(2) INFORMATION FOR SEQ ID NO: 62:

35

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1851 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

CACTAGTATA ATTTATAATT ATAACCTATT CTGATTTCCTT TTCAAATATT AGGTGTCCTA 60
GTGCTATATG AAGGTTTGCC ACTTCATCTT GCACTGTTCC CCAAACCTTG GACTGAGCTA 120
45 TGCCAGACTC AGTCTGCTAT GTCAAAAAC TGCATCAAGC TTTTGTGTGA AGATCCTGTT 180
TTCGCAGAAT ATATTAAATG TATCCTAATG GATGAAAGAA CTTTTTTTAA CAACAACATT 240
GTCTACACGT TCATGACACA TTTCTTCTA AAGGTTCAAA GTCAAGTGTT TTCTGAAGCA 300
50 AACTGTGCCA ATTTGATCAG CACTCTTATT ACAAACCTGA TAAGCCAGTA TCAGAACCTA 360
CAGTCTGATT TCTCCAACCG AGTTGAAATT TCCAAGCAA GTGCTTCTTT AAATGGGGAC 420
55 CTGAGGGCAC TCGCTTTGCT CCTGTCAGTA CACACTCCCA AACAGTTAAA CCCAGCTCTA 480
ATTCCAACCTC TGCAAGAGCT TTTAAGCAA TGCAGGACTT GTCTGCAACA GAGAACTCA 540
CTCCAAGAGC AAGAAGCCAA AGAAAGAAAA ACTAAAGATG ATGAAGGAGC AACTCCCAT 600

AAAAGGCGGC GTGTTAGCAG TGATGAGGAG CACACTGTAG ACAGCTGCAT CAGTGACATG 660
AAAACAGAAA CCAGGGAGGT CCTGACCCCA ACGAGCACTT CTGACAATGA GACCAGAGAC 720
5 TCCTCAATTA TTGATCCAGG AACTGAGCAA GATCTTCCTT CCCCTGAAAA TAGTTCTGTT 780
AAAGAATACC GAATGGAAGT TCCATCTTCG TTTTCAGAAG ACATGTCAAA TATCAGGTCA 840
CAGCATGCAG AAGAACAGTC CAACAATGGT AGATATGACG ATTGTAAAGA ATTTAAAGAC 900
10 CTCCACTGTT CCAAGGATTC TACCCTAGCC GAGGAAGAAT CTGAGTTCCC TTCTACTTCT 960
ATCTCTGCAG TTCTGTCTGA CTTAGCTGAC TTGAGAAGCT GTGATGGCCA AGCTTTGCCC 1020
15 TCCCAGGACC CTGAGGTTGC TTTATCTCTC AGTTGTGGCC ATTCCAGAGG ACTCTTTAGT 1080
CATATGCAGC AACATGACAT TTTAGATACC CTGTGTAGGA CCATTGAATC TACAATCCAT 1140
GTCGTCAAA GGATATCTGG CAAAGGAAAC CAAGCTGCTT CTGACATTA GGTGTAGCAT 1200
20 GTCTACTTTT AAGTCCCTCA CCCCCAACC CCATGCTGTT TGTATAAGTT TTGCTTATTT 1260
GTTTTGTGC TTCAGTTTGT CCAGTGCTCT CTGCTTGAAT GGCAAGATAG ATTTATAGGC 1320
25 TTAATTCTTG GTCAGGCAGA ACTCCAGATG AAAAAACIT GCATCTTCAG TATACTTCCT 1380
AAAGGCAAT CAGATAATGG ATATGTTTAA TGTAATTAAG AGTTCACITT AGTGGCTTTC 1440
ATTTAATATG GCTGTCTGGG AAGAACAGGG TTGCCTAGCC CTGTACAATG TAATTTAAAC 1500
30 TTACAGCATT TTTACTGTGT ATGATATGGT GTCCCTCTGTG CCAGTTTTGT ACCTTATAGA 1560
GGCAGATTGC CTCGATCGC TGTGGTTCTT ATTATCAAAA TTAAGTTTAC TTGTATACGG 1620
35 AACACCACA AGAAATTGA TTCTGTAAAG AATCCTCTTT AGCTGTGGCC TGGCAGTATA 1680
TAAATGGTGC TTTATTTAAC AGAATACCTG TGGAGGAAAT AAAGCACACT TGATGTAAAA 1740
ATAATGTTT TATTTTATT GACATGACTG ATTGATTGCT ATTCTGTGCA CTTAATTAAA 1800
40 CTGATTGTGA TGACTTWWAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA A 1851

45

(2) INFORMATION FOR SEQ ID NO: 63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3542 base pairs
50 (B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

55 TCCAATGCTG ATGAGCGTCT TCGCTGGCAG GCCAGCTCCT TGCCTGCTGA TGACCTTTGC 60
ACAGAAAATG CCATCATGCT GAAACGATTC AATAGGTATC CGCTGATCAT TGACCCCTCT 120
60 GGACAGGCCA CAGAATTCAT TATGAATGAA TATAAGGWTG GTAAGATCAC ACGGACCAGC 180

	TTCTCTGGATG ACGCCTTCAG AAAGAACTTA GAGAGTGCAC TGAGATTTCGG TAACCCCTT	240
5	CTGGTCCAGG ATGTGGAAAG CTACGATCCA GTTTTGAACC CCGTGCTGAA CCGTGAAGTG	300
	CGGCGAACAG GGGGGAGAGT GCTGATCACT CTCGGGGACC AGGACATAGA CCTGTGCGCA	360
	TCGTTTGTC TCTTCCTGTC CACCCGGGAT CCAACTGTGC AGTTCACACC AGATCTCTGT	420
10	TCCCGGGTTA CTTTGTGAAA CTCACAGTT ACCCGTAGCA GTTTACAAAG CCAGTGCTTA	480
	AATGAAGTAC TTAAAGCAGA AAGACCTGAT GTGGACGAGA AACGATCTGA TCTCTTAAA	540
15	CTTCAAGGGG AATTTCAGCT CCGTTTGGCT CAGCTGGAAA AATCTCTACT ACAAGCTCTG	600
	AACGAGGTGA AAGGGCGCAT TTTGGATGAC GACACGATCA TAACCACTCT GGAGAACCTG	660
	AAGAGAGAGG CTGCAGAGGT CACCAGGAAA GTTGAGGAGA CGGACATTGT CATGCAGGAG	720
20	GTGGAGACCG TGTCACGCA GTACCTCCCG CTCTCCACCG CCTGCAGCAG CATCTACTTC	780
	ACCATGGAGT CCCTCAAGCA GATACACTTC TTGTACCAGT ACTCCCTCCA GTTTTTCCTG	840
25	GACATTTATC ACAACGTCCT ATACGAGAAC CCGAACCTGA AGGGTGTAC CGACCACACA	900
	CAGCGCTGT CCATTATAAC AAAGGACCTC TTCCAGGTGG CGTTTAACCG AGTGGCTCGA	960
	GGCATGCTGC ATCAGGACCA CATTACCTTT GCCATGCTGC TGGCAAGAAT CAAACTGAAG	1020
30	GGCACCGTGG GGGAGCCAC CTACGATGCA GAATTCCAGC ACTTCTTGAG AGGAAATGAG	1080
	ATGTCTCTGA GTGCTGGCTC CACCCCAAG ATCCAGGGCC TGACTGTGGA GCAGGCGGAG	1140
35	GCGGTGGTGA GGCTGAGCTG CCTTCCCGCG TTAAAGGACT TGATTGCAAA GGTTCAGGCA	1200
	GACGAGCAAT TTGGCATCTG GCTGGACAGC AGCTCCCGG AGCAGACTGT GCCCTACCTC	1260
	TGGAGTGAAG AAACACCTGC AACACCCATT GGCCAGGCCA TCCACCGCTT GCTCTGATC	1320
40	CAGGCTTTCC GGGCCGATCG CCGTTGGCC ATGGCCACA GTTTGTGTTT AACAAACCTT	1380
	GGGAGTCTT TCATGTCCAT CATGGAGCAG CCGCTCGACC TGACCCACAT TGTGGSCACA	1440
45	GAGGTGAAGC CCAACACTCC TGTCTTAATG TGCTCTGTGC CTGGTTATGA TGCCAGTGA	1500
	CATGTCGAGG ACCTTGCAGC CGAGCAGAAC ACGCAGATCA CTTCAATTGC AATCGGCTCT	1560
	GCAGAAGGCT TTAACCAAGC AGATAAGGCA ATAAACACCG CTGTAAAGTC GGGCAGGTGG	1620
50	GTGATGCTGA AGAATGTGCA TCTGGCCCA GGGTGGCTGA TGCAGCTGGA GAAGAAGTTG	1680
	CATTCCCTGC AGCCGCATGC CTGCTTCCGA CTCTTCTCA CCATGGAGAT CAACCCCAAG	1740
55	GTGCCGTGTA ATCTGCTCCG TGCGGGCCGC ATCTTTGTGT TCGAGCCACC GCCAGGRTG	1800
	AAGGCCAACA TGCTGAGGAC GTTCAGCAGC ATTCCCGTCT CACGATATG CAAGTCTCCC	1860
	AACGAGCGTG CCCGCTGTGA CTTCTGCTG GCCTGGTTTC ATGCGATCAT CCAAGAACGC	1920
60	TTACGATACG CACCACTGGG GTGGTCAAAG AAGTATGAAT TTGGAGAGTC TGACCTGCGG	1980

	TCANYTTGCG ATACGGTGGA CACGTGGCTG GATGACACGG CCAAGGGCAG GCAGAACATC	2040
5	TCACCGGATA AGATCCCGTG GTCTGCACTA AAGACCTTAA TGGCCAGTC CATTTATGGC	2100
	GGGCGCGTGG ACAACGAGTT TGACCAGCGT CTGCTCAACA CCTTCCTGGA GCGCCTGTTT	2160
	ACAACCAGGA GTTTCGACAG TGAGTTTAAG CTGGCATGCA AGGTGACGG ACATAAAGAC	2220
10	ATTCAAATGC CAGATGGCAT GCAGGCGAGA GGAGTTTGTG CAGTGGGTGG AGTTGCTCCC	2280
	CGACACCCAG ACGCCCTCCT GGCTGGGCCT GCCCAACAAC GCGAGAGAG TCCTCCTTAC	2340
15	CACACAGGT GTGGACATGA TCAGTAAAT GCTGAAGATG CAGATGTTGG AGGATGAGGA	2400
	CGACCTGGCC TACGCAGAGA CTGAGAAGAA GACGAGGACA GACTCCACGT CCGACGGGCG	2460
	CCCTGCCTGG ATGCGGACAC TGCACACCAC CGCGTCCAAC TGGCTGCACC TCATCCCCCA	2520
20	GACGCTGAGC CACCTCAAGC GCACCGTGGA GAATATCAAG GATCCTTTGT TCAGGTTCTT	2580
	TGAGAGAGAA GTGAAGATGG GCGCAAAGCT GCTTCAGGAC GTTCGCCAGG ACCTTGACAG	2640
25	TGTCGTCCAG GTGTGCGAAG GAAAGAAGAA GCAGACCAAC TACTTGCGCA CGCTGATCAA	2700
	CGAGCTAGTG AAAGGGATCT TGCCTCGGAG CTGGTCCAC TACACGGTGC CTGCCGGCAT	2760
	GACCGTCATC CAGTGGGTGT CCGACTTCAG CGAGAGGATC AAACAGCTGC AGAACATCTC	2820
30	ACTGGCAGCT GCATCTGGTG GCGCCAAGGA GCTAAAGAAC ATCCACGTGT GCCTGGGTGG	2880
	CCTGTTGCTG CTGAGGCGT ACATCACTGC CACCAGGCAG TATGTGGCCC AGGCCAACAG	2940
35	CTGGTCCCTG GAGGAGCTCT GCCTGGAAGT CAACGTCACC ACCTCACAGG GCGCCACCCT	3000
	TGACGCTGTC AGCTTCGGAG TCACGGGTTT GAAACTTCAA GGGGCCACGT GCAACAACAA	3060
	CAAGCTGTCA CTGTCCAATG CCATCTCAAC CGCCCTTCCC CTGACGCAGC TGCGCTGGGT	3120
40	CAAGCAGACA AACACCGAGA AGAAGGCCAG TGTGGTAACC TTACCTGTCT ACCTGAACTT	3180
	CACCCGTGCA GACCTCATCT TCACCGTGGA CTTGAAATT GCTACAAAGG AGGATCCTCG	3240
45	CAGCTTCTAC GAGCGGGGTG TCGAGTCTT GTGCACAGAG TAAACTTTTC TAGCTGCCCC	3300
	TTTCTGTAAT AGTGAAAGTT GGTATTTAAC ATTTATTCAT TTTTAAATA TTTGGAAGGT	3360
	CTGAGCTTGT GAAAAGAAAG TGGTTGGTCT GAGGTGGAG GAAGCTGAAT GGAATCTGAC	3420
50	GGTGGGAGT GGTGGAAATT GGAAGGATAC CAGGAGGTAT TTGGGAAGGC CAATGGCGTG	3480
	GCTCCTTTGA GGAAATAAAA CACTAAGCAT GAAAAAAAAA AAAAACTTA CAANCCNCAA	3540
55	GG	3542

60 (2) INFORMATION FOR SEQ ID NO: 64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 883 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

10 AGGTGATTTT AATGATAGGT GTCATATATA GGACGGATAA TCTGTTTACA TTCTGTTCTT 60
CTCGATGCAC TCACAAGCGG GTAAC TAGGT GACAAGAAAA CAAAGATCTT ATTCAAAGA 120
GGTCTTACAG CAACCCAACG TCTCATCTTC CCATAGTAAA GATGACGGCG CCTTGAGGTA 180
15 AGCTACAGGC AACACCACTT CCGCGTTTCT CTTCGCCCTT GGTCCAAGAT GGCGGATGAA 240
GCCACGGGAC GTGTTGTGTC TGAGATCCCG GTGCTGAAGA CTAACGCCGG ACCCCGAGAT 300
CGTGAGTTGT GGGTGCAGCG ACTGAAGGAG GAATATCAGT CCCTTATCCG GTATGTGGAG 360
20 AACAACAAGA ATGCTGACAA CGATTGGTTC CGACTGGAGT CCAACAAGGA AGGAACTCGG 420
TGGTTTGGA AATGCTGGTA TATCCATGAC CTCCTGAAAT ATGAGTTTGA CATCGAGTTT 480
25 GACATTCCCTA TCACATATCC TACTACTGCC CCAGAAATTG CAGTTCCTGA GCTGGATGGA 540
AAGACAGCAA AGATGTACAG GGGTGGCAAA ATATGCCTGA CGGATCATTT CAAACCTTTG 600
TGGGGCCAGG AATGTGCCCA AATTGGACT AGCTCATCTC ATGGCTCTGG GGCTGGGTCC 660
30 ATGGSTGGCA GTGAAATCC CTGATCTGAT TCAGAAGGGC GTCATCCAAC ACAAGAGAA 720
ATGCAACCAA TGAAGATCA AGCCACTGAG GCAGGGCAGA GGGACCTTTG ATAGGCTACG 780
35 ATACTAWTTT CCTGTGCATC AACTTAACT CATCTAACTG TTCCCCGGAC ANCCTCCACT 840
CTAGTTGTTA CTAAGTANTG CAGTAGCATT NTGGGAAGA ACA 883

40

(2) INFORMATION FOR SEQ ID NO: 65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1541 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

45

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

GGCACGAGGT GGCTCTACC CTGGGCTCAT CTGGCTACAC AGGGACTCTA AACGCTTCCA 60
GATTCCCTGG AAACATGCCA CCCGGCATAG CCCTCAACAA GAAGAGGAAA ATACCATTTT 120
55 TAAGGCCTGG GCTGTAGAGA CAGGGAAGTA CCAGGAAGGG GTGGATGACC CTGACCCAGC 180
TAAATGGAAG GCCCAGCTGC GCTGTGCTCT CAATAAGAGC AGAGAATTCA ACCTGATGTA 240
60 TGATGGCACC AAGGAGGTGC CCATGAACCC AGTGAAGATA TATCAAGTGT GTGACATCCC 300

	TCAGCCCCAG GGCTCGATCA TTAACCCAGG ATCCACAGGG TCTGCTCCCT GGGATGAGAA	360
5	GGATAATGAT GTGGATGAAG AAGATGAGGA AGATGAGCTG GATCAGTCGC AGCACCATGT	420
	TCCCATCCAG GACACCTTCC CCTTCCTGAA CATCAATGGT TCTCCCATGG CGCCAGCCAG	480
	TGTGGGCAAT TGCAGTGTGG GCAACTGCAG CCCGGAGGCA GTGTGGCCCA AAACCTGAACC	540
10	CCTGGAGATG GAAGTACCCC AGGCACCTAT ACAGCCCTTC TATAGCTCTC CAGAACTGTG	600
	GATCAGCTCT CTCCCAATGA CTGACCTGGA CATCAAGTTT CAGTACCGTG GGAAGGAGTA	660
15	CGGGCAGACC ATGACCGTGA GCAACCCCTCA GGGCTGCCGA CTCTTCTATG GGGACCTGGG	720
	TCCCATGCCT GACCAGGAGG AGCTCTTTGG TCCCGTCAGN CTGGAGCAGG TCAAATTCCC	780
	AGGTCCTGAG CATATTACCA ATGAGAAGCA GAAGCTGTTC ACTAGCAAGC TGCTGGACGT	840
20	CATGGACAGA GGAATGATCC TGGAGGTCAG CGGTCATGCC ATTTATGCCA TCAGGCTGTG	900
	CCAGTGCAAG GTGTACTGGT CTGGGCCATG TGCCCCATCA CTTGTGTCTC CCAACCTGAT	960
25	TGAGAGACAA AAGAAGGTCA AGCTATTTTG TCTGGAAACA TTCCTTAGCG ATCTCATTGC	1020
	CCACCAGAAA GGACAGATAG AGAAGCAGCC ACCGTTTGAG ATCTACTTAT GCTTTGGGGA	1080
	AGAATGGCCA GATGGGAAAC CATTGGAAAG GAAACTCATC TTGGTTCAGG TCATTCCAGT	1140
30	AGTGGCTCGG ATGATCTACG AGATGTTTTC TGGTGATTTC ACACGATCCT TTGATAGTGG	1200
	CAGTGTCCGC CTGCAGATCT CAACCCCAAG CATCAAGGAT AACATCGTTG CTCAGCTGAA	1260
35	GCAGCTGTAC CGCATCCTTC AAACCCAGGA GAGCTGGCAG CCCATGCAGC CCACCCCCAG	1320
	CATGCAACTG CCCCTGCCG TGCTCCCCA GTAATTGTGA ATGCCATCTT CTTCCTTCTC	1380
	TTTTTTATAA TATGTACAT ATGGATTTTT TTATGTMTTA GATTTAAACA GCTTTTAAAT	1440
40	CTCTGTTTTT TGTGACAGTG TTAGAAGTTT GTGATTCTCC AAATATGCCT AGATTTAAAG	1500
	CTGATTTAAT TTATGGAAAA AAAAAAAAAA AAAAAAAAAA A	1541

45

(2) INFORMATION FOR SEQ ID NO: 66:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 732 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

50	AGAAAAATGAA TGTTAGAAGG TGCCTGCCGA GCGGGACAG AGTGTTTGCT CGCGCTGGAG	60
60	AAGGCTCTGC TCAGCCCTGA GAGTCCCTTC CTGCCCCACC GATACTGGCA CTTTAAAAAG	120

GAAGCTGACC GCACAGTGTC CAGACGAATT GGCCCCAGA AGATGGGGAG TTCTGTCTTG 180
CCCTTCTGTG TCTGCGTGAC CTCACCCAGC CTAGGAGGGA GGTGCATTCA GGGTAGATT 240
5 GCCTCTCATT CAAAGTTCTG GGGCTTTGGG CGGAAAACAG CCAGCTTTGG CGCTGTGGG 300
GAGACTCCTC CAGACCAGGA ACCCCAGAAG GAGACAGAGC CTGCCACATC CTCCCACGCC 360
AGGCCCTGGG CCAGGGTGAT TGGACTGAGA ATTTGGCCAC AACCAATTG ATGCTGGCTG 420
10 GAACCAGAGG CCAGAAAGCC TGGCCTTGTC CCCATGTGGG AGCCCTGTCC TCAGCCCTCT 480
TGTCCTCTTG AGCTCAGTGA ATTCCACCA GGTGCCACA GCTCCTGGAC TTCAAATTCT 540
15 ATATATTGAG AGAGTTGGAG AGTATATCAG AGATATTTTT GGAAAGGAGT TGGTCTATGC 600
AATGTCAGTT TGGAACTCTC TTGAAAGTTT AATGTTTTTA TTAGGAGATT TAAAGAAAAT 660
AAAGGTCTAC AATATCAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 720
20 AAAAAAAAAA AA 732

25

(2) INFORMATION FOR SEQ ID NO: 67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 629 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

35

TTAAGGAATT CGGCMGATC CCGGCAAGTA ACATGACTAA AAAGAAGCGG GAGAATCTGG 60
GCGTCGCTCT AGAGATCGAT GGGCTAGAGG AGAAGCTGTC CCAGTGTGGG AGAGACCTGG 120
40 AGGCCGTGAA CTCAGACTC CACAGCCGGG AGCTGAGCCC AGAGGCCAGG AGGTCCCTGG 180
AGAAGGAGAA AACAGCCTA ATGAACAAAG CCTCCAATA CGAGAAGGAA CTGAAGTTTC 240
TTCCGCAAGA GAACCGAAG AACATGCTGC TCTCTGTGGC CATCTTTATC CTCCTGACGC 300
45 TCGTCTATGC CTACTGGACC ATGTGAGCCT GGCACTTCCC CACAACCAGC ACAGGCTTCC 360
ACTTGCCCC TTGGTCAGGA TCAAGCAGG ACTTCAAGCC TCAATAGGAC CAAGGTGCTG 420
50 GGGTGTTCCT CTCCTAACCT AGTGTCAAG CATGGCTTCC TGGCGGCCCA GGCCTTGCCT 480
CCCTGGCCTG CTGGGGGGTT CCGGGTCTCC AGAAGGACAT GGTGCTGGTC CCTCCCTTAG 540
CCCAAGGGAG AGGCAATAAA GAACACAAAG CTGAAAAAAA AAAAAAAAAA AACTCGTAGG 600
55 GGGGGCCCGT ACCCAATCGC CCTTTCGTG 629

60

(2) INFORMATION FOR SEQ ID NO: 68:

(i) SEQUENCE CHARACTERISTICS:

5

- (A) LENGTH: 1751 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

10

CTGCTAGCCG GCCGGCGCAG GCTGCCGAGC GGGTGAGCGC GCAGGCCAGG CCAAAGCCCT 60
GGTACCGCG CGGTGCGGC CTCAGTCTGC GGCCATGGGG GCGTCCGCGC GGCTGCTGCG 120
15 AGCGGTGATC ATGGGGGCCC CGGCTCGGG CAAGGCACCC GTGTGCTCGC GCATCACTAC 180
ACACTTCGAG CTGAAGCACC TCTCCAGCGG GGACCTGCTC CGGGACAACA TGCTGCGGGG 240
CACAGAAATT GCGTGTTAG CCAAGGCTTT CATTGACCAA GGGAACTCA TCCCAGATGA 300
20 TGTCATGACT CGGCTGGCCC TTCATGAGCT GAAAAATCTC ACCCAGTATA GCTGGCTGTT 360
GGATGGTTTT CCAAGGACAC TTCCACAGGC AGAAGCCCTA GATAGAGCTT ATCAGATCGA 420
25 CACAGTGATT AACCTGAATG TGCCCTTTGA GGTCAATAAA CAACGCCTTA CTGCTCGCTG 480
GATTCATCCC GCCAGTGGCC GAGTCTATAA CATTGAATTC AACCTCCCA AAAGTGTGGG 540
CATTGATGAC CTGACTGGGG AGCCTCTCAT TCAGCGTGAG GATGATAAAC CAGAGACGGT 600
30 TATCAAGAGA CTAAAGGCTT ATGAAGACCA AACAAAGCCA GTCCTGGAAT ATTACCAGAA 660
AAAAGGGTG CTGGAACAT TCTCCGAAC AGAAACCAAC AAGATTTGGC CCTATGTATA 720
35 TGCTTTCTTA CAACTAAAG TTCCACAAAG AAGCCAGAAA GCTTCAGTTA CTCCATGAGG 780
AGAAATGTGT GTAACATTA ATAGTAAGAT GGGCAAACCT CCTAGTCCTT GCATTTAGAA 840
GCTGCTTTTC CTAAGACTTC TAGTATGTAT GAATTCCTTG AAAATTATAT TACTTTTATT 900
40 TCTACTGATT TTATTTTGA TACTAAGGAT GTGCCAAATG ATTCGGATAC TAAGATGCAT 960
CGTTTGAAAT CATCTAGTGT GTTGATGCA GTTATCCTCA AAAACATCAG CGATGTCTGA 1020
45 ACCTTTAAAA CATCTGTTAG AGCAAAATTA AAAGAGCATT TGGTAGTAAT CTAACTTTTT 1080
GTTCACTTAA TAAGTGGTTG ATAAAGTTTC CATATTTTTC TGGAAAAGTT AAAAAAGTT 1140
ACATGTCATT TGGAGAAAAT ACGTAATCAG AAATTTGTC ATAGATTGAT GCCAAAAAG 1200
50 ACATTTCCAG CATGTGGAA CATGGTGAGA CACTATATAA AATTCAGAA AGAAAGCAAC 1260
TGGATTACA GATTTATTGT GAGACACAAA TTCACTGCTG CCTTTACACT AAGAAATGTA 1320
55 TATGTTAACC ATATATGCTG TATTTATTTT GTCGTTAAGC ATACTTTCAG TTTACTCAGA 1380
ATTTTCAATT TGCTATAAAG ATGTATCAAT TAGCATATAG AAAAATATTA CTTTAAGATG 1440
ACTTGTTCC TTGAAAATA CCTGTGTAAT GAGGGTTATG ATTTGTGTCA AAAATGACA 1500
60

TAAGTGCTTT TACAAGCACC AAAGTTGAAT GAATTTTCAA CAAAATGTAA TTAAAGTCTA 1560
TGTTTTTCAGT TATGACTCAG GTTAAGAAAT GTGTTTTAGG ATCTACTTGC TGGTTTTTCT 1620
5 TTTTGATCCA AATGTGTGAT CTGCCCTGAT AAATAACAAG TTATNGTACC ATCTCCCCCG 1680
CCAATAAAAA AAAAAAAAAA AAAAAAAAAAC TCGAGGGGGG GCCCGGTACC CAATTCTCCG 1740
10 NAATAGGNAG T 1751

15 (2) INFORMATION FOR SEQ ID NO: 69:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 508 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:
GGCACGAGAT TATGTATTAA AATGTTTTTG AATGTGAAA TATTAGAATA TTGTTACTAT 60
25 TTGACCCAAC TCAAAATCTC CATGGGAAAA TACCTGTCGA TACCCACAGT ATTGTTGAAA 120
ATAATCAGAT GCAGTATCAC AGCTGTGTCA GACTCTAGTA CCAGTTGGGC AATCAAGGCA 180
30 CAGCTAAAAA TTGAAAACAA AGATCTGGAC AACAAAACAG CCAAAGGTGG GGGTCAAGAA 240
GCTCTGACGT GTACCTAGCT GTAGAATGCT ATGCACACGT GCCAGGTGTA GTGTGCATAT 300
CCAGGAAAAA CTGCAGAGAG CCCAGTCTT CACCTCTGGT TGACCATGAG CTCTGTGTAA 360
35 GCAGGAAGTG AAGGCTAAGG CAGATTTAAG CTCTGAAAGC ATTCCACAAC ATACACACAA 420
ATCGTGCAAA GCATTAAGGA AATCTTGTTA CTGCTAAGTG TTGCTGACCC AGGAACAAC 480
40 CCTACTCAGC TGGACTTAAA AATAAAAA 508

45 (2) INFORMATION FOR SEQ ID NO: 70:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 245 base pairs
(B) TYPE: nucleic acid
50 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:
55 TACATAGAGC AAAGAGAAAT TTCCAGAATT TCTARAATTC TGGAAAGAGA ATTTTCCTGA 60
GATTGCAGAT TTGCTTGTGT CCTCAGGTGA TGATGAGGCG TGTTCCTCCC TGTGTCTCTT 120
TCCTCACACT CATGCTTCCT CTCCTAGAGT GTCTGGTTGG CATGATCATG TGCTACCTAG 180
60

GCATTTCCTTT CACTGATACA AGGAAACTG CAGGGTTAAA AAAAAAAAAA AAAAAAAAAA 240
NCNCG 245

5

(2) INFORMATION FOR SEQ ID NO: 71:

10

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 361 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

ATGTTCTCTCA TGAGGATGCA CTGTGCTTC TGCAAGTATT GCTGCAGCTT CATAGTGACT 60
20 CCCACCAGCA CCAGCAATAC AGCTAGCTAC CTGTGGCCTT GGATCTCAGC CAGCATGGCT 120
GGGAGAGGGA GCAGCTGGGC ATGTACCCTA AATGCTGTTA CCAGGGAAGG ACTCCCAGAG 180
TGAAGACAAG TAGGGACTTC CTGCAGAGGT GGTACATGTG CTCTCTGTAT CCATACTTTT 240
25 TTTTCTTTT TTTTGTAGATA GAGTTTCACC CTGTGTGCCC TGGCTGGAGT GCAATGGTGC 300
GATCTCAGCT CACTGCAACC TCTCTGCCTC CCGGGTTCAA GTGATTCTCC TGCCTCAGCC 360
30 T 361

35

(2) INFORMATION FOR SEQ ID NO: 72:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 713 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

45 AGGATCACAC AATAGAGAAC ACTGTAGTAA CATTCGGTC TGCTCACAAG ACCCAGAACA 60
TTGATCAGTT TTTGTTGTTG GTTTATTATT TTTCTGTTAA AAAATTGTGA AAAGTTTGTT 120
TTAGCTAGAT GATATTTTAA TAGCTGCGAG TGCTTTGGAA CTATAAAGAT GTCACTACTT 180
50 AACACACATA CCTATGTTT TGTTTTGTMT TGTTTTACAC TCAGTATAAA TCAGGAGAAG 240
TTAGCCAACC ATCTAGCATT TAGAATCCTC TTTTATTATG TCTTCTAAGG ATATGGATGT 300
55 TCCCATAACA GCAACAAAAC AGCAACAAA ACATTTTATA AATATCACTT GATAGACTGT 360
AAGCACCTGC TTAACCTTGT GTCCCAAATA TTTAGTGTGT ATATATATAT ATATATATAC 420
ACACACACAC ACATATATAT TCAACAAATA AAGCAAATA TAACATGCAT TTCACATTTT 480
60

GTCTTTCCCT GTTACGATTT TAATAGCAGA ACTGTATGAC AAGTTTAGGT GATCCTAGCA 540
TATGTTAAAT TCAAATTAAT GTAAACAGA TTAACAACAA CAAAGAACT GTCTATTTGA 600
5 GTGAAGTCAT GCTTCTATT ATAATAACTT GGCTTCGGTT ATCCATCAA TGCACACTTA 660
TACTGTTATC TGATTGTTTA TAATAAGAA TACTGTACTT ATAAAAAAAA AAA 713

10

(2) INFORMATION FOR SEQ ID NO: 73:

(i) SEQUENCE CHARACTERISTICS:
15 (A) LENGTH: 862 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

GAAAGTCAGA GCTGTCCAAT CCCTCAGCAC CTTTATAGATT TGCTCCAAAT TAGAAACGTG 60
GGGACTATGT GTTCTGGGCA ATCAGAGTC TGGAAATGG CTCTGCAGGC TCTTGATAGT 120
25 GAGACAGTGG TCATCTTACC AGACATGCAT CTGATTTTAA GCCTCAGGCT AATCCACAAT 180
GCTCGGCCAT GCCTATGATT AACAAACAAA AGCAAAATCT GCTTTTATAG TTTAGGAAAC 240
30 CTGGATAGAA CAGTATTTTT CAGCATTTCT GGATAAGCA GTTCTGCATT TTAAATTGG 300
GACTGCAGAA GTGACTGTCT ATAGTTGTGA AATACAAAA ATGGTATGTT TGATCAGAAA 360
AGGAAGCCCG TGCTTGGCAC TTGGAAAGAT ACTGAGCATC ATAACCCTAA TGAGAAAATG 420
35 TAGGCTCTGT GAATGTTAAC TACAAATCAG GTTAGGAAAG CATATGACAC CCTTTGTCAA 480
ACTAAGCTTC ACTAGGAGGA CCTGTGCTCA TAGAAGAATA TGCTTTAAAA GTATCAATTT 540
40 TCCACAGTCG ATGATGGAGA AAAGTTCATT TGCACCAGAA TGCTGATAGT CACAATACAC 600
AGCCTGACAT ATATACAAT ACAGTTTCT GTAAACAGAA GTTCTTCCTC TTCCAATTCA 660
GGAGTCAGTC AGAGCATAAA TATTGCATGT TTCACTTTAG AAAGTATTC ATTTTAGAAA 720
45 GCAGATCTGG ATTATTTTGC AGGGTAGAAA TGAAGGCTAT TTCTGGCATT CTGCTCAA 780
AAGTCAATAT ATGTACATTA AGTATAAAAA AGGGTCTCTT TCACCTCTTT TGTTCGTAG 840
50 CATGGCTAC ATAACCTGTG CC 862

55 (2) INFORMATION FOR SEQ ID NO: 74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4602 base pairs
60 (B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

5	GCGAGGGGGC GKGGGGAGCA GCGCCGARGC CGCCGCCTCC GCCTCCGCCG CCTAGGACTA	60
	GGGGGTGGGG GACGGACAAG CCCCAGTGCC GGGGGAACG GAAGAGCCGA GACCCCCGGA	120
10	GCAGCAGGAC CAGGAAGGGG GAGAGGCGGC CAAGGCGGCT CCGGAGGACC CGCAACAACG	180
	GCCCCCTGAG GCGGTGCGCG CGGCGCCTGC AGGGACCACT AGCAGCCGCG TGCTGAGGGG	240
	AGGTGCGGAC CGAGGCCGGG CCGCTGCGRC CGCCGCGCMG CAGCTGTGTC CCGCCGAGGA	300
15	AGGCCGAGTA TCCCCGCCG CGAGGAGCAG CCCACGCGCC AGGCCTCCCG ACGTCCCCCG	360
	GCAGCAGCCC AGGCCGCGAA GTCCCCGTCT CCAGTTCAGG GCAAGAAGAG TCCGCGACTC	420
20	CTATGCATAG AAAAAGTAAC AACTGATAAA GATCCCAAGG AAGAAAAAGA GGAAGAAGAC	480
	GATTCTGCCC TCCCTCAGGA AGTTTCCATT GCTGCATCTA GACCTAGCCG GGGCTGGCGT	540
	AGTAGTAGGA CATCTGTTTC TCGCCATCGT GATACAGAGA ACACCCGAAG CTCTCGGTCC	600
25	AAGACCGGTT CATTGCAGCT CATTTGCAAG TCAGAACCAA ATACAGACCA ACTTGATTAT	660
	GATGTTGGAG AAGAGCATCA GTCTCCAGGT GGCATTAGTA GTGAAGAGGA AGAGGAGGAG	720
30	GAAGAAGAGA TGTTAATCAG TGAAGAGGAG ATACCATTCA AAGATGATCC AAGAGATGAG	780
	ACCTACAAC CCCACTTAGA AAGGGAAACC CCAAAGCCAC GGAGAAAATC AGGGAAGGTA	840
	AAAGAAGAGA AGGAGAAGAA GGAAATTAAA GTGGAAGTAG AGGTGGAGGT GAAAGAAGAG	900
35	GAGAATGAAA TTAGAGAGGA TGAGGAACCT CCAAGGAAGA GAGGAAGAAG ACGAAAAGAT	960
	GACAAAAGTC CACGTTTACC CAAAAGGAGA AAAAAGCCTC CAATCCAGTA TGTCCGTTGT	1020
40	GAGATGGAAG GATGTGGAAC TGTCTTGCC CATCCTCGCT ATTGTCAGCA CCACATTAAA	1080
	TACCAGCATT TGCTGAAGAA GAAATATGTA TGTCCCCATC CCTCCTGTGG ACGACTCTTC	1140
	AGGCTTCAGA AGCAACTTCT GCGACATGCC AAACATCATA CAGATCAAAG GGATTATATC	1200
45	TGTGAATATT GTGCTCGGC CTCAAGAGT TCCCACAATC TGGCAGTGCA CCGGATGATT	1260
	CACACTGGCG AGAAGCATTA CAATGTGAGA TCTGTGATT TACTTGTGCA CAAAAGGCAT	1320
50	CTCTTAATTG GCACATGAAG AAACATGATG CAGACTCCTT CTACCAGTTT TCTTGCAATA	1380
	TCTGTGGCAA AAAATTGAG AAGAAGGACA GCGTAGTGGC ACACAAGCA AAAAGCCACC	1440
	CTGAGGTGCT GATTGCAGAA GCTCTGGCTG CCAATGCAGG CGCCCTCATC ACCAGCACAG	1500
55	ATATCTTGGG CACTAACCCA GAGTCCCTGA CGCAGCCTTC AGATGGTCAG GGTCTTCCTC	1560
	TTCTTCCTGA GCCCTGGGA AACTCAACCT CTGGAGAGTG CCTACTGTTA GAAGCTGAAG	1620
60	GGATGTCAAA GTCATACTGC AGTGGGACGG AACGGGTGAG CCTGATGGCT GATGGGAAGA	1680

	TCCTTGTTGGG AAGCGGCAGC AGTGGAGGCA CTGAAGGGCT GGTATGAAC TCAGATATAC	1740
	TCGGTGCTAC CACAGAGGTT CTGATTGAAG ATTACAGACTC TGCCGGACCT TAGTGGACAG	1800
5	GAAGACTTGG GGCATGGGAC AGCTCAGACT TTGTATTTAA AAGTTAAAAA GGACAAAAAA	1860
	AAAATCTAAA GCATTTAAAA TCTAGTGAAA TAACTGAAGG GCCTGCTCTT TCCATTGTGG	1920
10	ATCACAGCAC ACACATACAT ACACCCTCCA CCTCCCCATC CCTGTCTCTC CCTCTGTTC	1980
	TCCCCTTATA AAATTGATGT TGTCCTTACC AGAAAGGTAG ACAAAAAAGA AGCAGCAGCA	2040
	GCTCTTAAAG TGAGGGTTAT TCTCATACTC GGTTCAGCC ATCAGCAGAC TTCCTGCTCA	2100
15	TCGGCAGATC CCCCTTTCCA ACCTGTAACCT CTGATGTGCT CTGGATCAGC TTTTAACCTT	2160
	TAATCATATA TTACTGTCTT CTAAATCCCT TCTCCTCCTC TACTGCTGCC CTATGGTTCT	2220
20	GGCTCCTACC CCTGCGGCA CACTTATCTT CAAATACCAT AGAATCTAA TCTCTGAAAT	2280
	CATAGCTCTC CAGTGGCTTT TAAAGAAAGC TGGTCCTCAG CACTAACAAA ATCACTACAA	2340
	TAGCCTAGTG CTTTTTTGGA AGCCTTTTTA GGAAGAATG TTAGGTTCAT GGTAACTAGT	2400
25	ATGCTCTTTG AGATTTTAC AGTGTGAAA CTTAAGAATT TTGAGAGGGT GAGGAGGGTT	2460
	GTTCAGAATC TAAATTACAG ATAGATGATT GTTCTTGTG AATTTGTTTC TTTTCCTTTT	2520
30	TTTTTGTCCT TACCATTTC TTACATTTC CTGGGGCCC ATCTCTGGCT CCTTGCTTTT	2580
	TGTTCTTGC TTTGCTTTAT CAGTTCATC CAGCTCCCTG TTAGTGAAGG AACTGCTGT	2640
	TAGTGAAGGA ACAAAGTCTA TGAGTCTAA AATTTTAAGT CAAAGAAAAC TGCTCTGTTT	2700
35	CCCCTTTAGT AACACTTCTG AAGAGGAAAA ACTTCAATAG CCAAAGTTAA TAATCCTATA	2760
	TAATAATTGC TTTGGCTTTC ACCTAAAATT CTGGGCATCA CAATTTCCTT GGGATAGAGG	2820
40	TTGTGTGGG GAATAGATTG CTTATTGCTG TTCACTGGAG AGAAAAGGTA GTGTTTTTGT	2880
	ACAAGGTCAT ACCGCCAGAA GCCCCAAATC CTATTTTGGC TCATCTTCAG GTAAAGAGTA	2940
	ATTCTATCC TGTGTGCTC AGAAGCTAGA ATCGAAGGCT TACCCTATTC ATGTATTATT	3000
45	GTCAGAAATG CATGATGGCT CTTGGAAAGA ATGACGTTTT GCTGGAAAAA AAAAAARAA	3060
	CMGTTTGTGT TTCACAAACA TGGCTTATCA ATTTTTTCAA AGAATCTTTT TTTCCCAAAA	3120
50	AGAGGAGTAA CAAAATGTCA TTTCTGAAAG AGGCTTACTT TATACCAACT AGTGTACGCA	3180
	TTTGGGATGC CAGGGAACAG AGAGTGAGAC ACCTACAATC ACCAGTCTCA AATGCGCTAT	3240
	TGTTCTTTTT CAGAGTGTG CAGATTTGCC ATTTCTCCAT AATATGGGGA TAGAAAATGG	3300
55	AATAAAGATA GAAGGGATGT AGAATATGCT TTCCTGCCAA CATGGTTTGG AGTCGACTTT	3360
	GGTATATTGA CTAGATTGTA AAATACAAGA TTGATTAGAT GAATCTACAA AAAAGTTGTC	3420
60	CTCCTCTCAG GTCCCTTTTA CACTTTTTGA CTAAGTAGCA TCTATATTCC AACTTAGCT	3480

TTTTTGTAC ACTTATCCTT TGTCTCCGTA AATTTCATTT GCAGTGGTTA GTCATCAGAT 3540
ATTTTAGCCA CCTACACAAA AGCAAACTGC ATTTTAAAA ATCTTTCTGA GATGGGAGAA 3600
5 AATGTATTCT CCTTTCCTAT ACCGCTCTCC CAACAAAAAA ACAACTAGTT AGTTCTACTA 3660
ATTAGAAACT TGCTGTACTT TTTCTTTTCT TTTAGGGGTC AAGGACCCCTC TTTATAGCTA 3720
CCATTTGCCT ACAATAAATT ATTGCAGCAG TTTGCAATAC TAAAAATTTT TTTATAGACT 3780
10 TTATATTTT CCTTTTGATA AAGGGATGCT GCATAGTAGA GTTGGTGTA TTAAGTATC 3840
TCAGCCGTTT CCTGCTTTC CCTCTGCTC CATATGCCCTC ATTGTCTTC CAGGGAGCTC 3900
15 TTTTAATCTT AAAGTTCTAC ATTTTCATGCT CTTAGTCAA TTCTGTTACC TTTTAAATAA 3960
CTCTTCCAC TGCATATTC CATCTGAAT TGGTGGTTCT AAATCTGAA ACTGTAGTTG 4020
AGATACAGCT ATTAAATATT TCTGGGAGAT GTGCATCCCT CTTCTTTGTG GTTGCCCAAG 4080
20 GTTGTMTTGC GTAAC TGAGA CTCCTTGATA TGCTTCAGAG AATTTAGGCA AACACTGGCC 4140
ATGGCCGTGG GAGTACTGGG AGTAAATAA AAATATCGAG GTATAGACTA GCATCCACAT 4200
25 AGAGCACTTG AACCTCCTTT GTACCTGTTT GGGGAAAAAG TATAATGAGT GTAATACCAA 4260
TCTAACTAAG ATTATTATAG TCTGGTTGTT TGAAATACCA TTTTCTCTC CTTTGTGTT 4320
TTTCCACTT TCCAATGTAC TCAAGAAAT TGAACAAATG TAATGGATCA ATTTAAAATA 4380
30 TTTTATTTCT TAAAGCCTT TTTGCTGTG TGTAATGTGC AGGACCCCTC TCCTTTCATG 4440
GGAGAGACAG GTAGTTACCT GAATATAGGT TGAAAAGGTT ATGTAAAAG AAATTATAAT 4500
35 AAAAGGATA CTTTGCTTTT CAAATCTTTG TTTTCTCTTA TTCTAGGTAA GGCATATTAA 4560
AAATAAATAT GTAAAGAAGA AAAATAAAG TTGTCTTCAT GG 4602

40

(2) INFORMATION FOR SEQ ID NO: 75:

45 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1255 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

CGCGCCCCGG GCCGGCGGGT TTCTCTAACA AATAAACAGA ACCCGCACTG CCCAGGCGAG 60
CGTTGCCACT TTCAAAGTGG TCCCCTGGGG GAGCTCAGCC TCATCCTGAT GATGCTGCCA 120
55 AGGCGCACTT TTTATTTTAA TTTATTTTAT ATTTTTTTAT TAGCATCCTT TTGGGGCTTC 180
ACTCTCAGAG CCAGTTTFTA AGGGACACCA GAGCCGACG CTGCTCTGAT TCTATGGCTT 240
60 GGTGTACT ATAAGAGTAA TTGCCTAACT TGATTTTCA TCTCTTAAAC CAACTTGTG 300

5 GCCAAAGAT ATTTGACCGT TTCCAAAATT CAGATTCTGC CTCTGCGGAT AAATATTTGC 360
CACGAATGAG TAACTCCTGT CACCACTCTG AAGGTCCAGA CAGAAGGTTT TGACACATTC 420
TTAGCACTGA ACTCCTCTGT GATCTAGGAT GATCTGTTCC CCCTCTGGAT GAACATCCTC 480
TGATGATCAA GGCTCCCAGC AGGCTACTTT GAAGGGAACA ATCAGATGCA AAAGCTCTTG 540
10 GGTGTTTATT TAAAATACTA GTGTCACTTT CTGAGTACCC GCCGCTTCAC AGGCTGAGTC 600
CAGGCCTGTG TGCTTTGTAG AGCCAGCTGC TTGCTCACAG CCACATTTCC ATTTGCATCA 660
TTACTGCCTT CACCTGCATA GTCACTCTTT TGATGCTGGG GAACCAAAAT GGTGATGATA 720
15 TATAGACTTT ATGTATAGCC ACAGTTCATC CCCAACCCTA GTCTTCGAAA TGTTAATATT 780
TGATAAATCT AGAAAATGCA TTCATACAAT TACAGAATTC AAATATTGCA AAAGGATGTG 840
20 TGTCTTTCTC CCCGAGCTCC CCTGTTCCCC TTCATTGAAA ACCACCACGG TGCCATCTCT 900
TGTGTATGCA GGGCTATGCA CCTGCAGGCA CGTGTGTATG CACTCCCCGC TTGTGTTTAC 960
ACAAGCTGTG GGGTGTACG CATGCCCTGCT TTTTTCACCT AATAATACAG CTTGGAGAGA 1020
25 TTTTGTATC ACATTATAAA TCCCACCTGC TCTTTTGTAT GGCCACATAA TAACTACTGC 1080
ATAATATGGA TACGCCTTAT TTGATTAAAC TAGTTCCTA ATGATGGACT TTTAAGTTGT 1140
30 TTCCTTTTTT TTTCTTTTTT GCTACTGCAA ACGATGCTAT AATAAATGTC CTTATCAAAA 1200
AAAAAAAAA AAAAAAAAAA AAAAAANCCC NGGGGGGGG CCCCGGGAAC NCAAT 1255

35

(2) INFORMATION FOR SEQ ID NO: 76:

40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 475 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:

GGCAGGAGAG AAATGTTTGA TTCTCTTTCC TATTTTAAGG GATCTTCTCT CTTGTTGATG 60
TTGAAAACCT ACCTTAGTGA AGATGTGTTT CAACATGCTG TTGTCCTTTA CCTGCATAAT 120
50 CACAGCTATG CATCTATTCA AAGTGATGAT CTGTGGGATA GTTTTAATGA GGTCACAAAC 180
CAAACACTAG ATGTAAAGAG AATGATGAAA ACCTGGACCC TGCAGAAAGG ATTTCCCTTTA 240
55 GTGACTGTTT AAAAGAAAGG AAAGGAACTT TTTATACAAC AAGAGAGATT CTTTTTAAAT 300
ATGAAGCCTG AAATTCAGCC TTCAGATACA AGGTACATGC CCTCTTCTTT TTCATGCCAT 360
CTCTTTTGCA CTCTCAGGTG GAAATATTTT GAAGTGTTTT ATAATCATAA GTTCTTGTGA 420
60

AACCTAACAA GATTATCCCT TCCTAAGAAT ACTTAACCTT CCTACCAAAT TAAAA

475

5

(2) INFORMATION FOR SEQ ID NO: 77:

(i) SEQUENCE CHARACTERISTICS:

10

(A) LENGTH: 465 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 77:

TTCTCTCTGC TCCTCGACTG CACCGCACTC GCGCGTGACC CTGACTCCCC CTAGTCAGCT 60

CAGCGGTGCT GCCATGGCGT GCGGCGGCG CGAACCRCG TCGGGGCTCG CGCGGTGTTG 120

20

GCTCTGGCGT TGCTCGCCCT GGCCTGTGC GTGCCCCGGG CCCGGGGCCG GGCTCTCGAG 180

TGGTTCTCGG CCGTGGTAAA CATCGAGTAC GTGGACCCGC AGACCAACCT GACGGTGTGG 240

25

AGCGTCTCGG AGAGTGGCCG CTTCGGCGAC AGCTCGCCCA AGGAGGGCGC GCATGGCCTG 300

GTGGGCGTCC CGTGGGCGCC CGGCGGAGAM CTCGARGGCT KCGCGCCCGA CACGCGCTTC 360

TTCTGTCCCG AGCCCGGCGG CCGAGGGGCC GCGCCCTGGG TCGCCCTGGT GGTCTGTTGG 420

30

GCTGCACCTT TCAAGGACAA AGTCTGGTG GCGGCGCNGA ANGAA 465

35

(2) INFORMATION FOR SEQ ID NO: 78:

(i) SEQUENCE CHARACTERISTICS:

40

(A) LENGTH: 1907 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 78:

45

ACATGCAGCC CAACTACAGA TTCTTATGGA ATTCTCAAG GTTGCAAGAA GAAATAAGAG 60

AGAGCAACTG GAACAGATCC AGAAGGAGCT AAGTGTTTTG GAAGAGGATA TTAAGAGAGT 120

50

GGAAGAAATG AGTGGCTTAT ACTCTCCTGT CAGTGAGGAT AGCACAGTGC CTCAAATTGA 180

AGCTCCTTCT CCATCACACA GTAGTATTAT TGATTCCACA GAATACAGCC AACCTCCAGG 240

TTTCAGTGGC AGTTCTCAGA CAAAGAAACA GCCTTGGTAT AATAGCACGT TAGCATCAAG 300

55

ACGAAAACGA CTTACTGCTC ATTTTGAAGA CTGGAGCAG TGTTACTTTT CTACAAGGAT 360

GTCTCGTATC TCAGATGACA GTCGAACTGC AAGCCAGTTG GATGAATTTC AGGAATGCTT 420

60

GTCCAAGTTT ACTCGATATA ATTCAGTACG ACCTTTAGCC ACATTGTCAT ATGCTAGTGA 480

TCTCTATAAT GGTTCAGTA TAGTCTCTAG TATTGAATTT GACCGGGATT GTGACTATTT 540
TGC GATTGCT GGAGTTACAA AGAAGATTAA AGTCTATGAA TATGACACTG TCATCCAGGA 600
5 TGCAGTGGAT ATTCATTACC CTGAGAATGA AATGACCTGC AATTTCGAAA TCAGCTGTAT 660
CAGTTGGAGT AGTTACCATA AGAACCTGTT AGCTAGCAGT GATTATGAAG GCACTGTTAT 720
10 TTTATGGGAT GGATTACAG GACAGAGGTC AAAGGTCTAT CAGGAGCATG AGAAGAGGTG 780
TTGGAGTGTT GACTTTAATT TGATGGATCC TAAACTCTTG GCTTCAGGTT CTGATGATGC 840
AAAAGTGAAG CTGTGGTCTA CCAATCTAGA CAACTCAGTG GCAAGCATTG AGGCAAAGGC 900
15 TAATGTGTGC TGTGTAAAT TCAGCCCCTC TTCCAGATAC CATTGGCTT TCGGCTGTGC 960
AGATCACTGT GTCCACTACT ATGATCTTCG TAACACTAAA CAGCCAATCA TGGTATTCAA 1020
AGGACACCGT AAAGCAGTCT CTTATGCAAA GTTTGTGAGT GGTGAGGAAA TTGTCTCTGC 1080
20 CTCAACAGAC AGTCAGCTAA AACTGTGGAA TGTAGGAAA CCATACTGCC TACGTTCCCTT 1140
CAAGGGTCAT ATCAATGAAA AAACTTTGT AGGCCTGGCT TCCAATGGAG ATTATATAGC 1200
25 TTGTGGAAGT GAAAATAACT CTCTCTACCT GFACTATAAA GGACTTTCTA AGACTTTGCT 1260
AACTTTTAAG TTTGATACAG TCAAAAGTGT TCTCGACAAA GACCGAAAAG AAGATGATAC 1320
AAATGAATTT GTTAGTGCTG TGTGCTGGAG GGCACCTACCA GATGGGGAGT CCAATGTGCT 1380
30 GATTGCTGCT AACAGTCAGG GTACAATTAA GGTGCTAGAA TTGGTATGAA GGGTTAACTC 1440
AAGTCAAATT GTACTTGATC CTGCTGAAAT ACATCTGCAG CTGACAATGA GAGAAGAAAC 1500
35 AGAAAATGTC ATGTGATGTC TCTCCCCAAA GTCATCATGG GTTTTGGATT TGTTTTGAAT 1560
ATTTTTTTCT TTTTTTCTTT TCCCTCCTTT ATGACCTTTG GGACATTGGG AATACCCAGC 1620
CAACTCTCCA CCATCAATGT AACTCCATGG ACATTGCTGC TCTTGGTGGT GTTATCTAAT 1680
40 TTTTGTGATA GGGAAACAAA TTCTTTTGAA TAAAATAAA TAACAAAACA ATAAAAGTTT 1740
ATTGAGCCAC AGTTGAGCTT GGAAAGTTT TGTCAAATGC NGCAAGAGAT AACTCTTTTT 1800
45 ANGAAGTAGC ATATGTGAAC TATAATGTAA CAGTGAATAA TTTGTAAAGT TCGTATTTCC 1860
CAACCTCTTT GGAATTACA CATATCAATA TAAACAAAAT ATAAAGT 1907

50

(2) INFORMATION FOR SEQ ID NO: 79:

(i) SEQUENCE CHARACTERISTICS:
55 (A) LENGTH: 1168 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

GCTGGGGTGT CCCCKCSGCC ACCATCGTCA TCGCTTACTT GATGAAGCAC ACTCGGATGA 60
5 CCCATGACTG ATGCTTATAA ATTTGTCAAA GGCAAACGAC CAATTATCTC CCCAAACCTT 120
AACTTCATGG GGCAGTTGCT AGAGTTCGAG GAAGACCTAA ACAACGGTGT GACACCGAGA 180
ATCCTTACAC CAAAGCTGAT GGGCGTGGAG ACGGTTGTGT GACAATGGTC TGGATGGAAA 240
10 GGATGCTGC TCTCCATTAG GAGACAATGA GGAAGGAGGA TGGATTCTGG TTTTITTTTCT 300
TTCTTTTTTT TTTTGTAGTT GGGAGTAAGT TTGTGAATGG AAACAACTT GTTTAAACAC 360
TTTTATTTTA ACAAGTGTA GAAGACTATA ACTTTTGATG CCATTGAGAT TCACCTCCCA 420
15 CAAACTGACA AATTAAGGAG GTTAAAGAAG TAATTTTTTT AAGCCAACAA TAAAAATATA 480
ATACAACTTG TTTCTCCCCC TTTTCTTTT AAGCTATTTG TAGAGTTTAT GACTAAATAG 540
20 TCTGTGCAGG TTCATAGACC GAAGATACTA CACACTTTAA ACCAATTAAA AAGAACCAAA 600
AGTAAATAGA AAAGACATTG AATCACCAAG GCCTGGGATC AACCTGGGCT GTCCACACAG 660
AAAACAAAAA CCCAACCAAA CCAAGCCCTG TTGTGCTCAC TGGTGCAAAG AGAAGATCAG 720
25 GGCAGCTTAA GTGGTCTAAG RATCCTTCAG GCATTCTTTA AGGAGAAAAA GGATACCTTT 780
GATTTGTGT GTTTCATGCT CTGGATTTT TTTTTTTTC CTCTCTGGG TTTAAGAGAT 840
30 TTTTTTTGAA ATAGTGAGGA ACTGACCATT ATATGCCTTC ACTGGCTTCT TGTGCAATAA 900
TATGATGTTT TAAGTGTGCA AACAAGTTAG AGCTGGCAGC TGAATGATAG ACAAATAGTG 960
CAAATTTGCC AGCTTGGAGA TAGAAAGGAA TTCAACAATA TATCAAATAC TTTCTTCCC 1020
35 ACCTTTTCC TTTTTTTTTT TTTTTTCTGA TTGATCTG GTTACAGTGC CATAAACCTT 1080
GTTACATATG TATATCAGAA TGTAAGAAAA AAAAATTTAT TTAATAATAT TTTTCGCAA 1140
40 AAAAAANNA AAAAATCGA GGGGGGCC 1168

45 (2) INFORMATION FOR SEQ ID NO: 80:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 1285 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 80:

55 AGAAAAATCAC ATCCTAACAA AGAAGTCTGT CTAAGACAGT ACATCTCCTG TTGAACTGTC 60
ATCTTTCCAC AGGACTTTCT GTTTTAGGG ATGAGACTAT TCTCTGCTTC ATCAAGGAAA 120
GAGAAATGTT CAGGGTTGTA GGGATGGCAC ACTTATTAGT TCTGCCTGTC TGAAAGGTTT 180
60

CTGCAGGACA GTTTGGTCAG AGCTGCAATT CTTAGTCCAT GGTCTAATGC TTGAGTATCT 240
 CTTCTTTCCC TTTCCTGTCT CAGGAATCAG CTGAGAATTC ATTCGATTGT CATGCCTCTA 300
 5 GCCCCTFACT GTGATTTGTT GGTGCACTT TCATTTGCTT TAGTTCTAGA ATCACCTGTT 360
 GACTCCTCAG ACTTCACCTA ACTTTGGAAA CTCTCTTTTG GAGGCTTCTC ATTTCCCCCT 420
 AATTCGTGTC TGCTGAGCC CTAGAATTTT CCCACCAACG AATTATTCCA GGTAGATCCT 480
 10 AAGTTGCTGG ATCTAGTTGA TATTTAAACA ATATCTAGTT GATATTTCTC ATTCAGTTGG 540
 ATCCAGAAAC CAGTATCTCT NAAAAACAAC CTCTCATACC TTGTGGACCT AATTTTGTGT 600
 15 GCGTGTGTGT GTGCGCGCAT ATGTATATAG ACAGGCACAT CTTTTTTACT TTTGTAAAAG 660
 CTTATGCCCTC TTTGGTATCT ATATCTGTGA AAGTTTAAAT GATCTGCCAT AATGTCTTGG 720
 GGACCTTGT CTTCTGTGTA AATGGTACTA GAGAAAACAC CTATATTATG AGTCAATCTA 780
 20 GTTGGTTTGA TTCGACATGA AGGAAATTTT CAGATAACAA CACTAACAAA CTCTCCCTTG 840
 ACTAGGGGGA CAAAGAAAAG CAAAACTGAC CATAAAAAAC AATTACCTGG TGAGAAGTTG 900
 25 CATAAACAGA ATTAGGTAGT ATATTGAAGA CAGCATCATT AAACAGTTAT GTTGTCTCC 960
 TTGCAAAAAA CATGTACTGA CTTCCTGTTG AGTAATGCCA AGTTGTTTTT TTTATTATAA 1020
 AACTTGCCCT TCATTACATG TTTCAAAGTG GTGTGGTGGG CCAAATATT GAAATGATGG 1080
 30 AACTGACTGA TAAAGCTGTA CAAATAAGCA GTGTGCCTAA CAAGCAACAC AGTAATGTTG 1140
 ACATGCITAA TTCACAAATG CTAATTTTCT TATAAATGT TTTGCTAAAA TACACTTTGA 1200
 35 AACTATTTTT CTGTATTCCA AGAGCTGAGA TCTTAGATTT TATGTAGTAT TAAGTGAAAA 1260
 AATACGAAAA TAATAACAT TGAAG 1285

40

(2) INFORMATION FOR SEQ ID NO: 81:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1290 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 81:

TCTCCAGCCC CAATTTCTAC GCGCACCAGA AGACGGAGGT CCTCTTTCCT TGCCTAACGC 60
 AGCCATGGCT CGTGGTCCCA AGAAGCATCT GAAGCGGGTG GCAGCTCCAA AGCATGATG 120
 55 GCTGGATAAA TTGACCGGTG TGTTTGCTCC TCGTCCATCC ACCGGTCCCC ACAAGTTGAG 180
 AGAGTGCTC CCCCTCATCA TTTTCCTGAG GAACAGACTT AAGTATGCCC TGACAGGAGA 240
 60 TGAAGTAAAG AAGATTTGCA TGCAGCGGTT CATTAAAATC GATGGCAAGG TCCGAACCTGA 300

TATAACCTAC CCTGCTGGAT TCATGGATGT CATCAGCATT GACAAGACGG GAGAGAATTT 360
CCGTCTGATC TATGACACCA AGGGTCGCTT TGCTGTACAT CGTATTACAC CTGAGGAGGC 420
5 CAAGTACAAG TTGTGCAAAG TGAGAAAGAT CTTTGTGGGC ACAAAGGAA TCCCTCATCT 480
GGTGACTCAT GATGCCCGCA CCATCCGCTA CCCCGATCCC CTCATCAAGG TGAATGATAC 540
10 CATTCAGATT GATTTAGAGA CTGGCAAGAT TACTGATTTC ATCAAGTTCC ATTACCCAG 600
CCAGGTGGTC TCGTCACCTC AGAGGCTCCG CAGACTCCTG CCCAGGCCAG GACTGAGGCA 660
AGCCTCAAGG CACTTCTAGG ACCTGCCTCT TCTACCAAG ATGAACTCAC TGGTTTCTTG 720
15 GCAGCTACTG CTTTCTCTCT GTGCCACCA CTTTGGGGAG CCATTAGAAA AGGTGGCCTC 780
TGTGGGAAT TCTAGACCA CAGGCCAGCA GCTAGAATCC CTGGGCCTCC TGGCCCCSGG 840
20 GGAGCAGAGC CTGCCGTGCA CCGAGAGGAA GCCAGCTGCT ACTGCCAGGC TGAGCCGTCG 900
GGGACCTCG CTGTCCCCGC CCCCGAGAG CTCGGGAGC CCCAGCAGC CGGGCCTGTC 960
CGCCCCCAC AGCCGCCAGA TCCCCGACC CCAGGGCGCG GTGCTGGTGC AGCGGGAGAA 1020
25 GGACCTGCCG AACTACAAC TGAACCTCTT CGGCCTGCGC TTCGGCAAGC GGGAGGCGGC 1080
ACCAGGAAC CACGGCAGAA GCGCTGGGCG GGGCTGAGGG CGCAGGTGCG GGGCAGTGAA 1140
30 CTTCAGACCC CAAAGGAGTC AGAGCATGCG GGGCGGGGCG GGGGGCGGG GACGTAGGGC 1200
TAAGGAGGG GCGCTGGAG CTTCCAACCC GAGGCAATAA AAGAAATGTT GCGTAACTCA 1260
35 AAAAAAAAAA AAAAAAANC TCGGGGGGG 1290

(2) INFORMATION FOR SEQ ID NO: 82:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 684 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 82:

TTTATTGTAT TCTGTAAC TAAGAACTCT ATTTWATTCT TTTTGGACT TGCTAAGTTG 60
50 TCTTTWATGG TTTTWAGTTC CATGCTGAAG TTTTCAGTAT TGACTTATCC CCTTGAACAT 120
GAGTTGTTTT ATAGACTCTR ATGATTCAAA AATCTTACAT CTTTGGTAG TCTCTTTCAT 180
55 TTGTYACTG TTTCTGTGA TTCTWACTCA TGGTATTTTA ATTCTTCGTT WTTTPTTTTC 240
TGTWAGAWA CATTCCTTGA AAAATAATTT GGAGGAATAT TTGATTCTTA TGAACAAGGC 300
ATTACTCACC AGAGAAGATT TTTTGTGTYT ACCARGTGCC TARGAATGCT AACAGTCTGG 360
60

GAMCACATAG AMCACCAGGT GATGAGACAA TCCTGGGART CCTGTTTTAC TTTGGSCCAT 420
CTTTTCTCCC AACCTGTGG GAATARTCAT YCATATCCTA RCTGCAGGCT ARAAGGTGGT 480
5 TTATCAGAGC CCAACTTCGA GGGCTCTGGG CTTTAGCTAC TGTACCCCA TCATAACTGA 540
GCTTCATGGA TTGATTCTCT TTTTATCTTT CAGATTTTCT TTTAAAAATC TTTGTTTTTT 600
TTTTTCTTCC GAAAGATTCC CCCAACATTA CCATTCCTCA CCTTCCTGTG AATTTTTTTG 660
10 GCTCTCATTT TGAATTTTTC AAGA 684

15

(2) INFORMATION FOR SEQ ID NO: 83:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2024 base pairs
20 (B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 83:

CTGCAGGAAT TCGGCACAGC TGGCTGGAG GCTTCATCTT TGCCGCCGCT GCCGTGCGCT 60
TCCTGGGATT GGAGTCTCGA GCTTCTCTCG TTCGTTCTGYC GCGGGGTTCTG CGCCCTTCTC 120
30 GCGCCTCGGG GCTCGGAGGC TGGGAAGGG GTTGGAGGG GCTGTTGATC GCCCGGTTTA 180
AGTTGCGCTC GGGGCGGCCA TGTGCGCCG CGAGGTCGAG CGCCTAGTGT CGGAGCTGAG 240
CGCGGGGACC GGAGGGGATG AGGAGGAAGA GTGGCTCTAT GGCGATGAAA ATGAAGTTGA 300
35 AAGGCCAGAA GAAGAAAATG CCAGTGCTAA TCCTCCATCT GGAATTGAAG ATGAAACTGC 360
TGAAAATGGT GTACCAAAAC CGAAAGTGAC TGAGACCGAA GATGATAGTG ATAGTGACAG 420
40 CGATGATGAT GAAGATGATG TTCATGTAC TATAGGAGAC ATTAAACGG GAGCACCACA 480
GTATGGGAGT TATGTTACAG CACCTGTAAA TCTTAACATC AAGACAGGG GAAGAGTTTA 540
TGGAACTACA GGGACAAAAG TCAAAGGAGT AGACCTTGAT GCACCTGGAA GCATTAATGG 600
45 AGTTCCACTC TTAGAGGTAG ATTGGATTG TTTTGAAGAT AAACCATGGC GTAAACCTGG 660
TGCTGATCTT TCTGATTATT TTAATTATGG GTTTAATGAA GATACCTGGA AAGCTTACTG 720
50 TGAAAAACAA AAGAGGATAC GAATGGGACT TGAAGTTATA CCAGTAACCT CTAATAACAA 780
TAAATTTACG GTACAGCAGG GAAGAACTGG AAACCTCAGG AAAGAACTG CCCTTCCATC 840
TACAAAAGCT GAGTTTACTT CTCCTCCTTC TTTGTTCAAG ACTGGGCTTC CACCGAGCAG 900
55 GAGATTACCT GGGCAATTG ATGTTATCG TCAGACTATA ACTATCAGCC GAGTAGAAGG 960
CAGGCGACGG GCAAAATGAGA ACAGCAACAT ACAGGTCCTT TCTGAAAGAT CTGCTACTGA 1020
60 AGTAGACAAC AATTTTAGCA AACCACCTCC GTTTTTCCTT CCAGGAGCTC CTCCCACTCA 1080

5
10
15
20
25
30
35

CCTTCCACCT CCTCCATTTC TTCCACCTCC TCCGACTGTC AGCACTGCTC CACCTCTGAT 1140
TCCACCACCG GGTTTTCCTC CTCCACCAGG CGCTCCACCT CCATCTCTTA TACCAACAAT 1200
AGAAAGTGGA CATTCCTCTG GTTATGATAG TCGTTCTGCA CGTGCATTTC CATATGGCAA 1260
TGTTCGCTTT CCCCATCTTC CTGGTCTGTC TCCTTCGTGG CCTAGTCTTG TGGACACCAG 1320
CAAGCAGTGG GACTATTATG CCAGAAGAGA GAAAGACCGA GATAGAGAGA GAGACAGAGA 1380
CAGAGAGCGA GACCGTGATC GGGACAGAGA AAGAGAACGC ACCAGAGAGA GAGAGAGGGA 1440
GCGTGATCAC AGTCCTACAC CAAGTGT TTTT CAACAGCGAT GAAGAACGAT ACAGATACAG 1500
GGAATATGCA GAAAGAGGTT ATGAGCGTCA CAGAGCAAGT CGAGAAAAAG AAGAACGACA 1560
TAGAGAAAGA CGACACAGGG AGAAAGAGGA AACCAGACAT AAGTCTTCTC GAAGTAATAG 1620
TAGACGTCGC CATGAAAGTG AAGAAGGAGA TAGTCACAGG AGACACAAAC AAAAAAATC 1680
TAAAAGAAGC AAAGAAGGAA AAGAAGCGGG CAGTGAGCCT GCCCCTGAAC AGGAGAGCAC 1740
CGAAGCTACA CCTGCAGAAT AGGCATGGTT TTGGCCTTTT GTGTATATTA GTACCAGAAG 1800
TAGATACTAT AAATCTTGTT ATTTTCTGCG ATAATGTTTA AGAAATTAC CTTAAATCTT 1860
GTTCTGT TTTG TTAGTATGAA AAGTTAACTT TT TTTCCAAA ATAAAAGAGT GAATTTTCA 1920
TGTTAAGTTA AAAATCTTTG TCTTGACTA TTTCAAAAAT AAAAAGACAG CAATGACTTT 1980
ATATCCAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAGGGC GGCC 2024

(2) INFORMATION FOR SEQ ID NO: 84:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 931 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 84:

40
50
55
60

CGCGCCMATA GCCGGACGGG GATCTGAGCT GGCAGGATGA ATGTGGGGGT GGCACACAGC 60
GAAGTAAACC CCAACACCCG AGTGATGAAT AGCCGAGGCA TCTGGCTGGC CTACATCATC 120
TTGGTAGGAT TGCTGCATAT GGTCTACTC AGCATCCCT TCTTCAGCAT TCCTGTTGTC 180
TGGACCTTGA CCAACGTCAT CCATAACCTG GCTACGTATG TCTTCCTTCA TACGGTGAAA 240
GGGACACCTT TTGAGACTCC TGACCAAGGA AAGGCTGGC TACTGACACA CTGGGAGCAA 300
ATGGACTATG GGCTCCAGTT TACCTCTTCC CGCAAGTTCC TCAGCATCTC TCCTATTGTG 360
CTCTATCTCC TGGCCAGCTT CTATACCAAG TATGATGCTG CGCACTTCCT CATCAACACA 420

GCCTCATTGC TAAGTGACT GCTGCCGAAG TTGCCCCAGT TCCATGGGGT TCGTGCTTT 480
GGCATCAACA AATACTGAGG GATGGGTTT GGGACAGCTC CATGGGCATG GGAAGGCAC 540
5 TGAAACAGAG GACTATAAAA CATCCTTCTC TTATTCTCCA TACTGTCTTC TACACCTTTA 600
AAGCCTGAGA ACTATACAAC CTTTCCCAGA CTCCAAGAA GAGAAGAGAT TGGCAAATGG 660
GGCTCCTGGG CCCAGTCTG CTAGTGGCAA GTTCTTTGA ATCAGGAAG CAGGTGAGGT 720
10 AAGGGCCAAA TCACTCTCCT CCATAGCAGG AAGCCATTG GGCAGCTCCT TTGGTGATTA 780
CATCTTTCCA TATCTTTTAC ACTTACCACC TTCCAGCTCT GTTTTGCTGT GTATTTTCT 840
15 TACAATAATT TTTTTCAGCT ATAGCTGCAG TTTAATCAGG ATGGGTAGAG AGCTGTCTC 900
ATAAGGCTGG GGGTGGGAAG ATGAATACT G 931

20

(2) INFORMATION FOR SEQ ID NO: 85:

(i) SEQUENCE CHARACTERISTICS:
25 (A) LENGTH: 825 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 85:

CGGGCCCGC GGGTCTTCA GGTACCGG CTGGTTACAG CAGCTCTACC CCTCACGACG 60
CAAACATGGC AGCGCAGAAG GACCAGCAGA AAGATGCCGA GCGGAAGGG CTGAGCGGCA 120
35 CGACCTGCT GCCGAAGCTG ATCCCTCCG GTGCAGGCCG GGAGTGGCTG GAGCGGCGCC 180
GGCGACCAT CGGCCCTGG AGCACCTTCG TGGACCAGCA GCGCTTCTCA CGGCCCGCA 240
40 ACCTGGGAGA GCTGTGCCAG CGCCTCGTAC GCAACGTGGA GTACTACCAG AGCAACTATG 300
TGTTCTGTGT CCTGGGCTC ATCCTGTACT GTGTGGTGAC GTCCCTATG TTGCTGGTGG 360
CTCTGGCTGT CTTTTTCGGC GCCTGTTACA TTCTCTATCT GCGCACCTTG GAGTCCAAGC 420
45 TTGTGCTCTT TGGCCGAGAG GTGAGCCAG CGCATCAGTA TGCTCTGGCT GGAGGCATCT 480
CCTTCCCTT CTCTGGCTG GCTGGTGCGG GCTCGGCCGT CTTCTGGGTG CTGGGAGCCA 540
50 CCTTGGTGGT CATCGGCTCC CACGCTGCCT TCCACCAGAT TGAGGCTGTG GACGGGGAGG 600
AGCTGCAGAT GGAACCGTG TGAGGTGTCT TCTGGGACCT GCCGCCTCC CGGGCCAGCT 660
GCCCCACCCC TGCCCATGCC TGCTCTGCAC GGCTCTGCTG CTCGGGCCCA CAGCGCGTC 720
55 CCATCACAAG CCCGGGAGG GATCCCGCCT TTGAAAATAA AGCTGTTATG GGTGTCATTC 780
AGGAAAAAAA AAAAAAAGG GGGGCCCTC TAGGGGTCAA AGTTA 825

60

(2) INFORMATION FOR SEQ ID NO: 86:

5

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1238 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 86:

CATGTAAAAG	GATGAAATGT	GACTTCTGGT	GTTTTTTTAT	TTCTATGGAG	GGACTTTCTG	60
GGGACGGTTT	CTGGCTCTCA	GGCTCTGAGA	AGCTGCAGTT	TATGAGTGGC	TCTGTGTGTG	120
CTGCCACCTA	CTGGAGAAGC	CATAAGCTGC	AGCTTTAGGA	AAAGGGAACC	CGGGGCAGAG	180
TGTGGGAAG	TGGATGGCA	GCATGGCAGG	GCTTTGGAAA	ATGAGAGGTG	AGAGTKTKTC	240
CAGGAAGGGT	GTAAGGAGAG	GATGGATCCT	GATACATGGA	TTCAGGATCA	TTAGGGTCCT	300
GTCTGGGACA	CTGGCCTTCC	TGCTTACCTG	CTCTTTCCTT	CCTCCTTGGT	CGGAGGAGGG	360
GCTGGCTCAC	TGCTCTGGCT	TCATTTTCCA	GAGCTGCCTG	CTGCAGTCAC	ACTTAGGTCA	420
TCTTCTCTCA	CTTTTCTCCT	TTTGCCGATT	AGTGGACGTG	ACAGAGATGT	GAATGGGGCA	480
GGGATGTCCT	TTGATGGCAT	CAAGACTTTA	GCTTCTGGTG	CGCTGTGTCC	CAGCTCTGAT	540
TTTCAGTTGCA	GCCGTGATGG	AMAGTTNGCA	TGGAAGCTGA	GACTCTCACT	GACAGTGAAA	600
CCCTCAAATG	AACACAATCC	CTGCTTTCCT	GCCAAGGATC	CTTGTAGGGT	NCCCCAGCT	660
TCCCCACTTT	TTTTCTGTGT	CCTGACAAAG	AAACACAGAG	TAACTTGATT	GCCCTGTGAC	720
CTGGCCAGTT	GCATTTCCCC	TGCAGGCTTG	AGCCCAAGCC	AGAGCCTTGA	AAAGGTATTC	780
AGGTGTGTGC	CCAAAACACT	GAAAAAACT	GCCCTGGCCC	TGAACCAAAT	ACCTTGAACC	840
CTCGTAAACT	CCATACCCTG	ACCCCCTTGT	TTTGGATATA	CCCAGGTAGA	ACAACTCTCT	900
CTCACTGTCT	GTGTGAGGA	TACGCTGTAG	CCCACTCATT	AAGTACATTC	TCCTAATAAA	960
TGCTTTGGAC	TGATCACCCCT	GCCAGTCTTT	TGCTCTGGGC	AATCTATACT	TTTNTCAGA	1020
GGTTCCCAAG	GCCTACTGAA	GGGACTTAAC	ATACTCTTAA	TGGCTTTCCT	CTCTCTTGTT	1080
TTACCTTATG	CCCTCACTTC	CTGAGTTAAC	CTCCCAAATA	CAGGATTCAC	CTGTACCCAA	1140
GCCCTTAGCT	TCAAGAATAC	AGGATCACCT	GTACCCAAGC	CCTTAGCTCA	AGCTCTGCTT	1200
TGGAAGAACC	CAAAC TAAGA	CAGTGCTCCT	GGTGCCCT			1238

55

(2) INFORMATION FOR SEQ ID NO: 87:

60

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1460 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 87:

	ATTGCCTTCT GGTCCCTGGT GACACTGGGG TCATCCTTCA TCCCCGGAGA GCATTTCCTGG	60
10	CTGCTCCTCC TGACCCGGGG CCTGGTGGGG GTCGGGGAGG CCAGTTATTC CACCATCGCG	120
	CCCACTCTCA TTGCCGACCT CTTTGTGGCC GACCAGCGCG ACCGGATGCT CAGCATCTTC	180
	TACTTTGCCA TTCCGGTGGG CAGTGGTCTG GGCTACATTG CAGGCTCCAA AGTGAAGGAT	240
15	ATGGCTGGAG ACTGGCACTG GGCTCTGAGG GTGACACCGG GTCTAGGAGT GGTGGCCGTT	300
	CTGCTGCTGT TCCTGGTAGT GCGGGAGCCG CCAAGGGGAG CCGTGGAGCG CCACTCAGAT	360
20	TTGCCACCCC TGAACCCAC CTCGTGGTGG GCAGATCTGA GGGCTCTGGC AAGAAATCCT	420
	AGTTTCGTCC TGTCTTCCCT GGGCTTCACT GCTGTGGCCT TTGTACGGG CTCCTGGCT	480
	CTGTGGGCTC CGGCATTCCCT GCTGCGTTCC CGCGTGGTCC TTGGGGAGAC CCCACCCTGC	540
25	CTTCCCGGAG ACTCCTGCTC TTCCTCTGAC AGTCTCATCT TTGGACTCAT CACCTGCCTG	600
	ACCGGAGTCC TGGGTGTGGG CCTGGGTGTG GAGATCAGCC GCCGGCTCCG CCACTCCAAC	660
30	CCCCGGGCTG ATCCCCTGGT CTGTGCCACT GGCTCCTGG GCTCTGCACC CTTCTCTTC	720
	CTGTCCCTTG CCTGCGCCCG TGCTAGCATC GTGGCCACTT ATATTTTCAT CTTTCATTGA	780
	GAGACCTTCC TGTCCATGAA CTGGGCCATC GTGGCCGACA TTCTGCTGTA CGTGGTGATC	840
35	CCTACCCGAC GCTCCACCGC CGAGGCCTTC CAGATCGTGC TGTCCACCT GCTGGGTGAT	900
	GCTGGGAGCC CCTACCTCAT TGGCTGATC TCTGACCGCC TGGCCCGGAA CTGGCCCCC	960
40	TCCTTCTTGT CCGAGTTCCG GGCTCTGCAG TTCTCGCTCA TGCTCTGCGC GTTTGTGTTGGG	1020
	GCACTGGGCG GCGCACTTCC TGGGCACCGC CATCTTCATT GAGGCCGACC GCCGGCGGGC	1080
	ACAGCTGCAC GTGCAGGGCC TGCTGCACGA AGCAGGGTCC ACAGACGACC GGATTGTGGT	1140
45	GCCCCAGCGG GGCCGCTCCA CCCGCGTGCC CGTGGCCAGT GTGCTCATCT GGAGAGGCTG	1200
	CCGCTCACCT ACCTGCACAT CTGCCACAGC TGGCCCTGGG CCCACCCAC GAAGGGCCTG	1260
50	GGCCTAAACC CCTTGGCCTG GCCCAGCTTC CAGAGGGACC CTGGGCCGTG TGCCAGCTCC	1320
	CAGACACTAC ATGGGTAGCT CAGGGGAGGA GGTGGGGGTC CAGGAGGGG ATCCCTCTCC	1380
	AACAGGGGCA GCCCAAGGG CTCGGTGCTA TTGTAAACGG GATTAAAATT TGTAGCCAGA	1440
55	AAAAAAAAA AAAAAAAAAA	1460

60

(2) INFORMATION FOR SEQ ID NO: 88:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 1395 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 88:

10 CAGGTGCAAA GTGGGAAGTG TGAGTCCTCA GTCTTGGGCT ATTCCGGCCAC GTGCCTGCCG 60
GACATGGGAC GCTGGAGGGT CAGCAGCGTG GAGTCCTGGC CTTTTCGCTC CACGGGTGGG 120
15 AAATTGGCCA TTGCCACGGC GGGAACTGGG ACTCAGGCTG CCCCCCGGCC GTTTCCTCATC 180
CGTCCACCGG AYTCTGTGGG GCTCGCACTG GCGCTGATGT AGTTTCCTGA CCTCTGACCC 240
GTATTGTCTC CAGATTAAAG GTACGACATT TGGAGGCCCC AGCGAGAAAC GTCACCGGGA 300
20 GAAACGTCAC CGGGCGAGAG CGGKCCCGCT GTGTGCTCCC CCGGAAGGAC AGCCAGCTTG 360
TAGGGGGGAG TGCCACCTGA AAAAAAATT TCCAGGTCCC CAAAGGGTGA CCGTCTTCCG 420
25 GAGACAGCGG ATCGACTACC ATGTGGGTGC CCACAAAAAT TYCACCTYTG AGTCTCAAC 480
TGCTGACCCC GGGGTCACTT CCAGAGAGAA GGACTCCCTC CTGCTTGGAA GAGACCTCAC 540
ACCGTCATCA CGATGCCAAC GGCTCTGAAG GTGGATGGCA TTCCTGCGTG GATTCATCAC 600
30 TCCCGCATCA AAAAGGCCAA CRGAGCCCAA CTAGAAACAT GGGTCCCCAG GGCTGGGTCA 660
GGCCCCCTAA AACTGCACCT AAGTTGGGTG AAGCCATTAG ATTAATTCTT TTTCTTAATT 720
35 TTGTAAACA ATGCATAGCT TCTGTCAACT TATGTATCTT AAGACTCAAT ATAACCCCT 780
TGTTATAACT GAGGGAATCA ATGATTGAT TCCCCAAAA CACAAGTGGG GAATGTAGTG 840
TCCAACCTGG TTTTACTAA CCCTGTTTTT AGACTYTCCT TTTCTTTAA TCACTCAGCC 900
40 TTGTTCCAC CTGAATGAC TCTCCCTTAG CTAAGAGCGC CAGATGGACT CCATCTTGGC 960
TCTTTCNACT GGCAGCCGCT TCCTYCAAGG ACTTAACTTG TGCAAGCTGA CTCCCAGCAC 1020
45 ATCCAAGAAT GCAATTAACT GATAAGATAC TGTGGCAAGC TATATCCGCA GTTCCCAGGA 1080
ATTCGTCCAA TTGATTACAC CMAAAGCCC CGCGTCTATC ACCTTGTAAT AATCTTAAAG 1140
CCCCGTCACC TGGAACTATT AACGTTCTTG TAACCATTTA TCCTTTTAA TTTTTCCT 1200
50 ACTTTATTTT TGTAAATTG TTTTAACTAG ACCCCCCCTC TCCTTTCTAA ACCAAAGTAT 1260
AAAAGCAAAT CTAGCCCTTT CTTCAGGCCG AGAGAAATTC GAGCGTTAGC CGTCTCTTGG 1320
55 CCACCAGCTA AATAACGGA TTCTTCATGT GTAAAAAAA AAAAAAAA CTCGGAGGGG 1380
GGCCCCGGTA CCCAA 1395

(2) INFORMATION FOR SEQ ID NO: 89:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 1186 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 89:

GGCACGAGCC GGCAAGCCGA GCTAGGGTGA AAAC TGGGGG CGCACCAGGA TGTNNGACAG 60
AAAAGCAGAA GATGAGACTC TGTTCATTCA CTTTTCCTAG GCCCATCCTG TGGTCATCTT 120
15 TCCCCCTCCC ATCATACCTC CTCCTTCCTG GAGCCTCTGC CGGCTTGGCT GTAATGGTGG 180
CACTTACCTG GATATTTTCA TGGGAGGATG AAAGGCGAGA CTCACCCTAC GCGGTGGGAC 240
20 AGATGGGGAG AGGAAAAAGG CAGAGATGGC CAGGAGAGGG GTGCAGGACA AACCAGAGAG 300
GTTGGGTGAG GGGAAAAAGG TGGGAGAGAA GAGGGGTGCA GGCCCTGCAG GCCGGTTAGC 360
CAGCAGCTGC GGCCTCCCCG GGCCTTTGGC ATCCAAC TTC GCAGACAGG TACCAGCCTC 420
25 CTGGTGTGTA TCATAGGATT TGTTCACATA GTGTATGCA TGATCTTCGT AAGGTTAAGA 480
AGCGGTGGTG GTGCACCATG ACATCCAACC CGTATATATA AAGATAAATA TATATATATA 540
30 TGTATGTAAA TTATGGCACG AGAAATTATA GCACTGAGGG CCCTGCTGCC CTGCTGGACC 600
AAGCAAAACT AAGCCTTTTG GTTTGGGTAT TATGTTTCGT TTTGTTATTT GTTTGTTTTT 660
GTGGCTTGTC TTATGTCGTG ATAGCACAAG TGCCAGTCGG ATTGCTCTGT ATTACAGAAT 720
35 AGTGTTTTTA ATTATCAAT GTTCTAGTTA ATGTCTACCT CAGCACCTCC TCTTAGCCTA 780
ATTTTAGGAG GTTGCCCAAT TTTGTTTCTT CAATTTTACT GGTACTTTTT TTGTACAAAT 840
40 CAATCTCTTT CTCCTTTCT CTCCTCCCA CCTCTCACC TTGCCCTCTC CATCTCCCTC 900
TCCCGCCCTC CCTCCTCCC TCTGGCTCCC CGTCTCATTT CTGTCCACTC CATTCTCTCT 960
CCCTCTCTCC TGCTCTCTGC TGCCCCCTCC CCAGCCCACT TCCCGAGTT GTGCTTGCCG 1020
45 CTCCTTATCT GTTCTAGTTC CGAAGCAGTT TCACTCGAAG TTGTGCAGTC CTGGTTCAG 1080
CTTCCGCAT CTGCCTTCGT TTCGTGTAGA TTGACGCGTT TCTTTGTAAT TTCAGTGT TT 1140
50 CTGACAAGAT TTAACAAAAA AAAAGGAAA AAAAAA AAAA 1186

55 (2) INFORMATION FOR SEQ ID NO: 90:

(i) SEQUENCE CHARACTERISTICS:

- 60 (A) LENGTH: 1821 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 90:

5	AAAACATGCT TTCAGGGCGT CCCCTATGTA TTCGGGGGGC CCACGGACAC TCAGGCTGGA	60
	KATCOGTCCT CACTGCGCTC AAGATGGCCT CAGCAGACAC CAGTTACCCA GCTGAAAGTC	120
	ACAAATCCCTC CCAGAAGTCT CCCAACACTA GTGCTGACCA GAGGTGGGGC TCTCAGGCTA	180
10	GGAGTTTCAC ACACAATGAC AGGCTGCTGG GGGACATTGC AGGACCCCTT TTCCTYTCCT	240
	CTCCATGCTA GAAGCCAGCC CTAGGMAGCT GCAGTTACTC CCTGTGACTC AGCAGCAGGC	300
15	TGATTCAACA CAGCTGCCCC CACAAAGCCA GTGGTAATAC ATCTGTTTAC CTTTCCCTAT	360
	CACCCAGACA CAAGCCCTT TCCCAGGTCA AACCACAGGC CGATGCATCT CCAGTTTGAC	420
	AGTCAAATCA CTAATTCCAT TGCTACTTTA GATCAGCCAA AGTGGTGACT GCTGCAGTGT	480
20	GTGGCTATCC CTACAAGGCC CACCAAGGG ATGCCCAAAG CCAACCTTC TCCAGGGCTG	540
	CAGCAGNAGC AACCCACCA GCCTAAGTCC AGCAGAGGAC CTCCCACCCA ATGTCTTGTT	600
25	CTAATTAGAA GGGGAAGTTA GCCACAGAAA ATCAACTTAT CTATAATTAC AAAATTCTCT	660
	TGACTCACCT TAAAGTTCCT ATTGACATCT ACTGCTTTTA AACCTATTG AAAACTCTGA	720
	TACTAAAACA AATGACACTC TAAGAAAGTT TGGGAGCCCC ATGCTGAGAA CCATTTCTGT	780
30	GCAGTGAGGA TGTTCACAGA AGCTACTTAC CTACATGTGA ATGTGCCATT TTCTTTCCTT	840
	TTGTAGAGAA AATCCCTTT ACTTTTGGGA ACAGTAATGG CAGCTTCTAG TACAGCCATT	900
35	ACAGTTTCAT ATGAGAAAAA TTAAGAATAA CTATAAAAT GTTAAATAT CCAATAATGG	960
	ATAATGATGG CCAGAAGATT TAACATACAA AGTAATTCTC AATGTAAAGC TATTCAGCTC	1020
	TTCCAGGTG AATGCCCTGT AACCCACCT GACCTTCCAC ATCATCTTCA AAAAGCAGTT	1080
40	TCTCTGTCC CCATGATTCT CCTATAAGGT AACTCTTTAG TCCTCCATTT AGCACATTTT	1140
	AAATCCTCCA AAGAATAAGT ATCATGTGAT TATTTTAGCT TTACAAAAA AAAGTTGAAT	1200
45	GGCGTTTTAT TTTCATGGCC TATAAGCAGG TACCTTAGTA GGCAGATAT AGGAAAAACA	1260
	AATTAGAGCA AAACAAATCC TCTACAAATC CAAGGCAGGA AAAGTGGTGG CAGAGTGAAT	1320
	CATTCTCCTG TCCCTCCCAT CAGGTCAAAT CAGGAGGCTG CAGTGAATGC CTGTTCCTTG	1380
50	AATGTGTAGC AGTTGTTCTT GTAACCTTTT AAAACTTGCC TATAGGCTGT TTAGCACAGT	1440
	ACAGATTAAA GATACAGTTA CGTAAACAGC AAAGTAATTT TATAGTGCTT CATCCATTTA	1500
55	TCATGCTTTG GTTTGCTAAT TTTTTCACAT ACCTTTTCT ATCAGTCT GTTGCTTTTG	1560
	TACACATTTT TCATATTGGG GTTCGACAGG TAAACACAAA CTGCTATTTT AGTAGAAAAA	1620
60	GTTATGTGTA TGGAATATTA AACCCAATAA ATGTATATAA GGGTAAAAA AAAAAAAAAA	1680

AAAAAAAAA AAAAAAAAAA AAAAAAATTC CTGCGGGCCG CANGCTTTT CCCTTTGGGT 1740
GAGGGGTAT TTTNGGCTTG GGCACGGGC CCTTCGTTT TACAACGTCG TGANGGGGG 1800
5 AACCGGGGG GGGTTTCCCC C 1821

10 (2) INFORMATION FOR SEQ ID NO: 91:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 862 base pairs
(B) TYPE: nucleic acid
15 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:

20 TGCCCTTTT CCCACCGATT CGGGCCTGG TGAAGGTGGG AGATGTGAAC TCCAATTAG 60
GGACTGGAGA GAGGTGAAGA ATTTGTCAGG TGGGAGATT GGATTGAAT GTGGACTTGT 120
AAATGACTTG ACCTTGCCAT CTGTGTCAA GGTACGGTT TGCTGTGGG TTCCTGGGAG 180
25 AGCTTACTCA CCCCGAGTC TTTCTTTCT CTGTCTCAA GAAGAGCCCT GTTGGTGCTT 240
TACCACCGCT TGGAGTCTCC CGAGGACACA AACAGGCAGA GAGGACCTG TAGGGAGAGT 300
30 TCTTCCTGT TTTCTGTGCT TTCTTTTTA CAGGACTCCC GGAAGGCCAC TCATGGCCAT 360
GCCAGGAGCT TTCTCAGAAA CAGTCATAAA CGATCTCTTG AGTCTCTTTC TTGTCTCTCC 420
AGCTGAGCTT TCTTATTCCA CCCTTTCTGG TGTCTATAGG AATGCATGAG AAGACCCTGG 480
35 GACGTTTTTC TGCTCTCTC TGGCCCTCCA TGGAGCCATG GGCCTCGGCC TCGGCGGCTC 540
CTCACCTCA CAATTTATTT CCTCTCCCG TGCCAGCCCT TCTTTTGTGT CTGAAACCGG 600
40 TTTTAAATG TGA CTCTCCC AGAGAAGAAG CCGCTGGCTG TATGAACTT GACGGCGCTT 660
TTGTAAGGTG CCACCCCAA ACTTTAAGGT AGCTAAACCA ATTTTAAAA GATTCAATGG 720
CTGTTCATC CTCCAGATGT AGCTATTGAT GTACACTTCG CAACGGAGTG TCTGAAATTG 780
45 TGGTGGTCCT GATTATAGG ATTCATAAT TAAATGTCT GCTGAATAAA AAAAAAAAAA 840
AAAAACTCGA GGGGGGCCG GT 862

50

(2) INFORMATION FOR SEQ ID NO: 92:

55 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 696 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
60 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 92:

CTGAGGCGAG TGAAGTGGAC TCTGAGGGCT ACCGCTACCG CCACTGCTGC GGCAGGGGCG 60
5 TGGAGGGCAG AGGGCCGCGG AGGCCGAGT TGCAAACATG GCTCAGAGCA GAGACGGCGG 120
AAACCCGTTC GCCGAGCCCA GCGAGCTTGA CAACCCCTTT CAGGACCCAG CTGTGATCCA 180
GCACCGACCC AGCCGGCAGT ATGCCACGCT TGACGTCTAC AACCTTTTGT AGACCCGGGA 240
10 GCCACCACCA GCCTATGAGC CTCCAGCCCC TGCCCCATTG CCTCCACCTT CAGCTCCCTC 300
CTTGAGCCC TCGAGAAAGC TCAGCCCCAC AGAACCTAAG AACTATGGCT CATACAGCAC 360
15 TCAGGCCTCA GCTGCAGCAG CCACAGCTGA GCTGCTGAAG AACAGGAGG AGCTCAACCG 420
GAAGCAGAG GAGTTGGACC GAAGGAGCGA GAGCTGCAGC ATGCTGCCCT GGGRGGCACA 480
GCTACTCGAC AGAACAATG GCCCCTCTA CCTTCTTTT GTCCAGTTCA GCCCTGCTTT 540
20 TTCCAGGACA TCTCCATGGA GATCCCCAA GAATTTTCTA AGACTGTATC CACCATGTAC 600
TACCTCTGGA TGTGCAGCAC GSTGGNTCTT CTCTGAAYT TCMTCGSCTG CCTGGCCAGT 660
25 TCTGTGTGGA AACCAACAAT GCGAGGCTT TGGGT 696

30 (2) INFORMATION FOR SEQ ID NO: 93:

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 1886 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 93:

40 CAGGCCACTG ACGCTTCTTT GCGAGGGATG CAGGAGGTCC TACAGAGAAA GCGCTTCTT 60
GCATKTCAGA GGGCCACAG CCTGTCACCC ACAGATCACC AAGCAGCTTT CTACCTGGCT 120
CTGCAGCTTG CCATCTCCAG ACAGATCCCA GAGGCTCTGG GGTATGTCCG CCAAGCTCTT 180
45 CAGCTTCAAG GTGACGATGC CAACTCCCTG CACCTCCTTG CCTCCTGCT GTCAGCACAG 240
AAGCATTACC ATGACGCTCT GAACATCATC GACATGGCCC TGAGTGAATA CCCAGAAAAT 300
50 TTCATACTAC TGTTTTCCAA AGTGAAGTTG CAGTCACTCT GCGAGGCCC GGACGARGCA 360
CTGCTGACTT GTAAGCACAT GCTGCAGATA TGGAAATCCT GCTACACCT CACCAACCCC 420
AGTGATTCTG GACGTGGGAG CAGCCTCTTA GATAGAACCA TTGCTGACAG ACGACAGCTT 480
55 AATACAATTA CTTTGCCAGA CTTGAGCGAT CCCGAGACAG GCTCCGTCCA TGCCACATCG 540
GTAGCAGCCT CAAGAGTGGA GCAGGCACTG TCGGAAGTGG CTTCGTCTCT GCAGAGCATG 600
60 CCCCTAAGCA GGGCCGCTG CACCCTGGA TGACGCTGGC ACAGATCTGG CTCCATGCAG 660

CTGAAGTCTA TATCGGCATC GGAAGCCTG CAGAAGCCAC AGCCTGTACC CAAGAAGCTG 720
5 CCAACCTCTT CCCAATGTCC CACAATGTCC TCTACATGCG CGGCCAGATT GCTGAGCTCC 780
GGGGAAGCAT GGACGAGGCG CGGCGGTGGT ATGAAGAGGC CTTAGCCANT CAGCCCCACC 840
CACGTGAAGA GCATGCAGCG ACTTGGCCCT GATCCTTCAC CAGYTAGGCC GYTACAGTYT 900
10 GGCGGAGAAG ATCTCTCGGG ACGCGGTGCA GGTGAACTCG ACAGCCCACG AGGTCTGGAA 960
CGGCTGGGC GAGGTCTCC AAGCTCAGGG CAACGATGCG GCGGTACGG AGTGCTTCCT 1020
GACAGCCTTG GAGCTGGAGG CCAGCAGCCC CGCCGTGCCC TTCACCATCA TCCCCCGGT 1080
15 GCTCTGAGCA GGCGCCTGCC AGCCTCACCT GCCGCTCAGC CTNCAGAGGC CCTGCCGGGC 1140
ACCAGGCTT GTGCCATCG CCCAAGGGA TGAATCTGCC GCACTGAGGC CAGGGACGAG 1200
20 TGTCAGTGG GCCACAGTA ACCAACCAA CCAACCCGA ATCATCGTC TCGCCATGTG 1260
CGTTCTCTT GTTTTTTTT CCAGCCCAAT GGTAGTTCT GAACCTATTG ACATTGTTCA 1320
AAATGGATCA TGTGCCATAT TTTGTTAGT GACATCTGAG TTTTCAGTAA AATGATTATG 1380
25 GAATTAATCA GCAATGTAG AAGAATATAT TCAAAGTTAA AATTCAGTGG CAGCACAGAT 1440
TATTTTTATC AGAGCTGTAA AGAAAACAAC TGTCTTTTC TCCCCACCAC CCCTCCTGCC 1500
30 CCACCTTGGC CCAGAAACCA AATGTGAAT TCCTGTCTCC CACCTCAGCA CTAGTCCATG 1560
CCAGGACACC AGCTGACAAT TTCTTGGTTT TACTGTCAAT AATGTACCA TGTGATCAAT 1620
TACTGTCTC ACTTAGAACA AAGCCTGAGT CCGAGAATAT TTATATTTTA CCAATATATG 1680
35 CCTGTTACAA GAGAAGGAAA TATGAGTTAT TTAAGTTTAA CTTTTTTATG TGAATTCAGA 1740
GTTTATTTAT CGAGGGAAAT ATGTACAAAG AAGCTTCAA TGGAAATTTT ACCGACATTC 1800
40 CTTATACATG ACAGACACTT GGCTACATGG GAAGATGATG TTAATAATAA AATGATTTTT 1860
AAATGGAAAA AAAAAAAAAA AAAAAN 1886

45

(2) INFORMATION FOR SEQ ID NO: 94:

50 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1774 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:

CTCAGCTACC GTATACAGTA GGACATAACC CCATTTTACA TGCACTACAC TGAGACTTGC 60
60 CTCCTCTCCC CCCACATTGA AGATGTTCTT TTTTCATAAC TATATACTAT TCCATTGCAT 120

	GAATATTCCTG TAATTTATTT AATCCCCTAT GGATTGATAA TTAGGTTTCAT TATAGATAGA	180
	AGTGTAATTA ACATTCTCTGT ACATGTATTT TGCTACTTGT GTGGGTATTT CTGTAGGATG	240
5	AATAACTAGA AATTTATTGG ATCAGGTTTC ACATTTCAG TTTTGAAAAC TACTACCAAA	300
	AAGATTTCAC CAATTTACAA CTCCATCATT AGTAAGAATG CCTGTTTGCC TATAGTCTGC	360
10	CAACCCTGAA TCCTTAAAAA TTTTGGCCAA TCTGGTAGGC AAAATTTCTT TCTTTTCTTT	420
	GAATATTAAT GAGGAGGAAC ATCTTTTCAT GTTCTTGCC CATTTCATT TCCTATTATG	480
	AATTGCTTTT GCCCATTTTC CTTTTTTTAA TTATGAAAGT CTAATGACTA CCTTCTCATT	540
15	GTATAAAAAA CACAGTTCTT TGAATAGAGA GACCCTTTTC TCCAATGCTA CCAATCACAT	600
	TCCACTTACC ACAGTTTAAC ATACATCCTC TAGTCACCTT TCCGTACGAA TATACATACA	660
20	CATAAAAAACA CTTTTTACAT AAATAGGATC TCATATTCTG TAGCTTTTAA AAATTTGGT	720
	CTCAAAAAAA GATAACAGGT CTTTAAATTT CTTTAATGGT TGAATATGAT TAAATACTAT	780
	GAAAATGCCA TTATTTATTTC CCTTAATTTT TTTCTCTCG CTATTACATT GCCAAAGTAA	840
25	ACATCCTATT CAGATGTCTT TGTGCATGTG TGTGAATATT TCTTTAGTCT GGAGTCCAGT	900
	AAGGTGGATT TTTGGATCAA AGGGTTTGTT CTCTGTCCAC CTTCAGTCTT CCCAAAGGCC	960
30	TTCATAACTG TATTTTCACC AAGTGTATGG AGAATGTTCA TTTCCCCATA TAACCATACC	1020
	TACACTTGAT AGTTTTTATC TGTGGGCGA AAAAGAACCT TTTCTTATTT TGCATTTCCC	1080
	TGATTATAAA AAAAAATGGT GAGATTGGGG TTATTTTCAT GTTTATTGGC CATTTATAGT	1140
35	TTACTGTGGA TTGTTTGTAT CCCTTACCTG CTTTCTATTG GGTATGTGT GGATATATTG	1200
	TTTTTATTTG TTCAGCATCT CCTTCCCAT CTTCTGGTAA CACAACCTTT ATTTATTTGT	1260
40	GGGGAACCTA TTCCCTGTGG CTTAGGTGAG CATGTGACCA GGCCTGGCCT CCTGAGTCCC	1320
	ACAGCTTCCT AGCCACAGTG ATAAAAGAAT GGGTATATAA CTTAAGCCAG GCTAAGGAAA	1380
	GCCCTTAACA GAACTTCTGC TGGAACTACT GGAAGAAGG CTTTATGGAG ATCCCAGGAA	1440
45	CCAAGGACCA TGTAAGCCTG AATTTGTGCC ATGTGGAGAG AGTCTGTCTG AGGAGAAACT	1500
	CGGATGCTAG CAGAAATGGA AAGAGAACTA AGTTCTGATG TCATTTTCTT GGAGGCCCTA	1560
50	GATCCAGCTG TGCCTAAAGC CTGCCCTACT CCGGACTTTA AAGTTTGTG AGCCAATAAA	1620
	GTCCCTTTCT TGTTAAGAT AATTGAATTG AGTTTCTGTT CTGATTAATA TAGGTTATTT	1680
	GTATTTTCTT ATTGATTGTG AGAAAACCTT TGTAATTTTA AATTCTAGAC TTTATGCACT	1740
55	ATATAAGTTA ATAAAATTAG CATGGCCTTC CATG	1774

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2503 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 95:

10	GGCACGAGCG AAGCAAGGG GGCACCAGCT CAGGACTGCA TCTGCCTGCC ATTTCCCTTC	60
	CACTCCTCCT TTCTGGAGTC TGACATTAGA AAGCCAGCGA GAAGGAAGAT TCAAACAACC	120
	AACCCTGATT TCCTGCTTCT CCTTTTCATG AGTGTTCCTG TGGTCTCTGC ACCTCCTTTC	180
15	TGTCCCCCGG CAGAGGGCAG TAGAGATGGC CGGCCAAGG CCTCRGTGGC GCGACCAGCT	240
	GCTGTTTCATG AGCATCATAG TCCTCGTGAT TGTGGTCATC TGCCTGATGT TATACGCTCT	300
20	TCTCTGGGAG GCTGGCAACC TCACTGACCT GCCCAACCTG AGAATCGGCT TCTATAACTT	360
	CTGCCTGTGG AATGAGGACA CCAGCACCTT ACAGTGTAC CAGTTCCTG AGCTGGAAGC	420
	CCTGGGGGTG CCTCGGGTGG GCCTGGGCCT GGCCAGGCTT GCGGTGTACG GGTCCCTGGT	480
25	CCTCACCTTC TTTGCCCCC AGCCTCTCCT CCTAGCCCAG TGCAACAKTG ATGAGAGAGC	540
	GTGGCGSCTG GCAGTGGGCT TCCTGGCTGT KTCTCTGTG CTGCTGGCAG GCGGCCTGGG	600
30	CCTCTTCCTC TCCTATGTGT GGAATGGGTC ARGCTCTCCC TCCCGGGGCC TGGGTTTCTA	660
	GCTCTGGGCA GCGCCAGSC CTTACTCATC CTCTTGCTTA TAGCCATGGC TGTGTTCCCT	720
	CTGAGGGCTG AGAGGGCTGA GAGCAAGCTT GAGAGCTGCT AAAGGCTTAC GTGATTGCAA	780
35	GGGTTCAATT CCAACCATGG TCAGAGGTGG CACATCTGCT CAGCCATCTC ATTTTACAGC	840
	TAACGCTGAT CTCAGCTCC AGCGATGGAA CCCACTACAG AGGAGGTGGG GCCCCTGTGT	900
40	CAAAGAGGCC GAGGGGCAGC AAGGGCAGMC AGGGCACCTG TGACTTCTTA GTACAAGATT	960
	GTCTGTCTTT CAGGACTTCC AAGGCTCCCA AAGACTCCCT AAACCATGCA GCTCATTTGT	1020
	ACACCAATTG CTGCTTTAAT TAATGGATCT GAGCAAATCT TCCTCTAGCT TCAGGAGGGT	1080
45	GGGAGGGGAG TGATTGCTGT CATGGGGCCA GACTTCCAGG CTGATTTGCC AAATGCCAAA	1140
	ATGAAACCTA GCAAAGAACT TACGGCAACA AACGAGGACA TTAAAAGAGC GAGCACCTCA	1200
50	GTGTCTCTGG GGACATGGTT AAGGAGCTTC CACTCAGCCC ACCATAGTGA GTGGGCCGCC	1260
	ATAAGCCATC ACTGGAATC CAACCCAGAG GGTCCAGGAG TGATCTCTGA GTGACTCAAC	1320
	AAAGACAGGA CACATGGGGT ACAAAGACAA GGCTTGACTG CTTCAAAGCT TCCCTGGACC	1380
55	TGAAGCCAGA CAGGGCAGAG GCGTCCGCTG ACAAATCACT CCCATGATGA GACCCTGGAG	1440
	GACTCCAAAT CCTCGCTGTG AACAGGACTG GACGGTTGCG CACAAACAAA CGCTGCCACC	1500
60	CTCCACTTCC CAACCCAGAA CTTGGAAGA CATTAGCACA ACTTACGCAT TGGGAATTG	1560

5
10
15
20
25
30
35

TGTTGATTTT CTAGCACTTG TGTATTGGAA AACCTGTATG GCAGTGATTT ATTCATATAT 1620
TCCTGTCCAA AGCCACACTG AAAACAGAGG CAGAGACATG TACTCTGGTG TGATCTCTTG 1680
TCCTCAGTGT CTCCTCTGGG CTCCTGTCCC TCTTGCTTTA TAGCTAGCTG CCCGGGGACC 1740
AAGGTACAGG TGAAAGCAAG GTAGCAGCTT GCGGGAGGAG GCCTGTCTGG CTTACCAGTC 1800
TATACACTGT GGCCTCAACC TCCCAGACAG GGCAGAGAAC TGTGGGCAGC TCGTTTGCTT 1860
TCTAGGCTGG CTGGAGAGGT GGGAGCTCAT TGATAGACTC ATGATGGAAA CTATTTTGA 1920
AACAGGCTTC CTCCTTCAGG AGAGATCATG CGGACTAAAC TGTAGCAATT CCAGTGCAAC 1980
TGGCAGTGAT CCTTTTCTTT GCAAAGTACT GTCTCTTTGG TTCCAGTAAG TTGGACCACC 2040
ACATGACATY ATTTTCCCTG GAACCTGGTC ACTGACTAAC ACAGACAATT GGGACTCCAG 2100
AGCCTCAAGA GCCAGGAGAG GGCACAGTAC ATACAGAGGG AGTCAAATGG GATCTCATTT 2160
TGAGTCTGTC CTTCGCGACA CTCAGAACGG CANCCCCAAG GCCCGGAGTG TCCAGGGCTT 2220
CTGGCCTGAG GTGAATCTGC CAGGCCCAAG AAGGCACAAA GGTAGGAGCA CAGAGAGCCC 2280
CATTTCCACA GCGGKCGGC CCAGCAGCAC CAGTGGAAGC TCAGCTGTCC TCCAGCTGCT 2340
CTCGGCAGAC AGTTCAGTGC ACAGTTTATG CCCTAGCTGA AAAAGATCTC CCGGACGTAT 2400
TTCAGCACAT CCTCTTCTC CTCCTCTCA GGGCTCCTGC TACAGGCAGA GCTGGAACCC 2460
CCCGGCCTCT GGAAGGGCT GAGGCCTGGA GYCAGTGCCT GTC 2503

(2) INFORMATION FOR SEQ ID NO: 96:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2801 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:

40
50
55
60

CTGGAAAGCC GAGGGTAGCC GAGCGGGGCG GCGCTCTGG AGCGGCGGGT GCTCGGGCTG 60
CCGTCCGCTC CGCCAGAAGC ACCGAGCAGC CGAGCCGGGG CCCGCGCCCC TCCTCCTCCA 120
TGAGGCCCGA GTGAGGCGCG GCGGCTATAG CCGACCCGCG GCGCCTTCCC CCCGCTCTCT 180
ATCGCGAGCG CACGACMAGC GGCCCTTGGA GGAGGAGGCG GAGGAGGAGG AGCATGTCCG 240
ACGGTTTCGA TCGGGCCCCA GGTGCTGGTC GGGGCCGGAR CCGGGCCTG GGCCGCGGAG 300
GGGCGGGGCC TRAGGGCGGC GGTTTYCCGA AMGGARCGGR GCCTGCTGAG CGGRCGCGGC 360
ACCAGCCGCC GCAACCCAAA GCCCCGGGCT TYCTGCARCC AMCGCCGCTG CGCCARCCCA 420

	GGACGACCCC GCCGCCAGGG GCCCAGTGCG AGGTCCCCGC CAGCCCCCAG CGGCCTTCCC	480
	GGCCCGGGGC GCTCCAGAG CAAACGAGGC CCCTGAGAGC TCCACCTAGT TCACAGGATA	540
5	AAATCCACACA GCAGAACTCG GAGTCAGCAA TGGCTAAGCC CCAGGTGGTT GTAGCTCCTG	600
	TATTAATGTC TAAGCTGTCT GTGAATGCCC CTGAATTTTA CCCTTCAGGT TATTCTTCCA	660
	GTTACACAGA ATCCTATGAG GATGGTTGTG AGGATTATCC TACTCTATCA GAATATGTTT	720
10	AGGATTTTTT GAATCATCTT ACAGAGCAGC CTGGCAGTTT TGAAACTGAA ATTGAACAGT	780
	TTGCAGAGAC CCTGAATGGT TGTGTTACAA CAGATGATGC TTTGCAAGAA CTGTGGAAC	840
15	TCATCTATCA ACAGGCCACA TCTATCCCAA ATTTCTCTTA TATGGGAGCT CGCCTGTGTA	900
	ATTACCTGTC CCATCATCTG ACAATTAGCC CACAGAGTGG CAACTTCCGC CAATTGCTAC	960
	TTCAAAGATG TCGGACTGAA TATGAAGTTA AAGATCAAGC TGCAAAAGGG GATGAAGTTA	1020
20	CTCGAAAACG ATTTTCATGCA TTTGTACTCT TTCTGGGAGA ACTTTATCTT AACCTGGAGA	1080
	TCAAGGGAAC AAATGGACAG GTTACAAGAG CAGATATTCT TCAGGTTGGT CTTGAGAAT	1140
25	TGCTGAATGC CCTGTTTTCT AATCCTATGG ATGACAATTT AATTTGTGCA GTAAAATTGT	1200
	TAAAGTTGAC AGGATCAGTT TTGGAAGATG CTTGGAAGGA AAAAGGAAAG ATGGATATGG	1260
	AAGAAATTAT TCAGAGAATT GAAAACGTTG TCCTAGATGC AAAGTGCAGT AGAGATGTAA	1320
30	AACAGATGCT CTTGAAGCTT GTAGAATCC GGTCAAGTAA CTGGGGCAGA GTCCATGCAA	1380
	CTTCAACATA TAGAGAAGCA ACACCAGAAA ATGATCCTAA CTACTTTATG AATGAACCAA	1440
35	CATTTTATAC ATCTGATGGT GTTCCCTTCA CTGCAGCTGA TCCAGATTAC CAAGAGAAAT	1500
	ACCAAGAATT ACTTGAAAGA GAGGACTTTT TTCCAGATTA TGAAGAAAAT GGAACAGATT	1560
	TATCCGGGGC TGGTGATCCA TACTTGATG ATATTGATGA TGAGATGGAC CCAGAGATAG	1620
40	AAGAAGCTTA TGAAAAGTTT TGTTTGGAAT CAGAGCGTAA GCGAAAACAG TAAAGTTAAA	1680
	TTTCAGCATA TCAGTTTAT AAAGCAGTTT AGGTATGGTG ATTTAGCAGA ACACAAGAGA	1740
45	GCAAGAAAAT GTGTCACATC TATACCAAAT TRAGGATGTT GAGTTATGTT ACTAATGTAT	1800
	GCAACTTTAA TTTTGTTAA CACTATCTGC CAAAATAAAC TTTATTCCT ATAACTTAAA	1860
	ATGTGTATAT ATATATAATA GTTTATTATG TACAGTTAAT TCTACTGTTT TGGCTGCAAT	1920
50	AAAATCGATT TTGAAATAAA TGAAATGTTG AAAATTTTGC TAGTTGGTTA GATGCTTATC	1980
	CTTTAAATTC TACTTTTCTT GAGGGGAAAA AGTCTTCGTC TGGAAATACA TATTACTGCA	2040
55	AAAATGTAGC ATCCTTTTTT AGGTAGGAGT ATTATAGCTT YCATTTTAGT TKGACATTTA	2100
	GTGTCCCAAT GAATTGAATT TCAAATATGA ATCATAATCT TGAAAATCTT TAGCACTAAA	2160
60	GTCTTGGGAA TATATCAACA ACTGATTTAC ATATGCAGAT GCTATTTGNA TACCAAGGGC	2220

TTTTTAAATG TCATGGGGG GAAAAACCCA ACTTGGTGGA ACTCCCAGCT AAACAACCAA 2280
GACTTCACTG GAAGATTAT TCCAATTCTA GGAATTGTC TTTTATTAT TTATTTTTC 2340
5 AACTGRCTAA CTTCATTACC TTAAAGCCTA GAACATTATT CTGCTTTATT TATATGGCTT 2400
TCTCACTTTT ATTTGTAGC AKGGGTGCA TCGACTTTT TACTAGAGAA TTTTACTAGA 2460
TATTTGTCAT TCAAGTTTC ATCTGCTTTA TAATTGATAC ACCTTGAGG TCACTTTTCT 2520
10 AATACTTTTA CTATAATGTG GTACCACCTC AGCCCTAATA AATAATATT TTACCTAATG 2580
TCAAATCTTT TTCCAGCTAA CTAAAACTG TGTACAAAAG GATTGCTTGT AAATATGCAT 2640
15 GTAAATAGTT CTGTTAATAA CCCACTGTTT TACATTGGT ACATCTGTGT CTGCTAATAC 2700
AGTTAGCTTT CTCACTTTC TGCTTGTTG TTCAGTCTGA ATTAAATTA GACTTTGAAA 2760
ATAAAGCTTA AAAAAAAAAA AAAAAAAAAA AAAAAGCTGA G 2801
20

(2) INFORMATION FOR SEQ ID NO: 97:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1631 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

30

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 97:

ATGGAGCCAA AGACAATCAC TGATGCTTTG GCTTCTAGTA TAATTAAGAG TGTGCTGCCT 60
35 AATTTTCTTC CATACAATGT CATGCTCTAC AGTGATGCTC CAGTGAGTGA ACTGTCCCTC 120
GAGCTGCTTC TGCTTCAGGT TGTCTTGCCA GCATTACTCG AACAGGGACA CACGAGGCAG 180
40 TGGCTGAAGG GGCTGGTGG AGCGTGGACT GTGACCGCCG GATACTTGCT GGATCTTCAT 240
TCTTATTTAT TGGGAGACCA GGAAGAAAAT GAAAACAGTG CAAATCAACA AGTTAACAAT 300
AATCAGCATG CTCGAAATAA CAACGCTATT CCTGTGGTGG GAGAAGGCCT TCATGCAGCC 360
45 CACCAAGCCA TACTCCAGCA GGGAGGCCT GTTGGYTTTC AGCYTTACCG CCGACCTTTA 420
AATTTTCCAC TCAGGATATT TCTGTTGATT GTCTTCATGT GTATAACATT ACTGATTGCC 480
50 AGCCTCATCT GCCTTACTTT ACCAGTATTT GCTGGCCGTT GGTAAATGTC GTTTTGACG 540
GGGACTGCCA AAATCCATGA GCTCTACACA GCTGCTTGTG GTCTCTATGT TTGCTGGCTA 600
ACCATAAGGG CTGTGACGGT GATGGTGGCA TGGATGCCTC AGGACGCAG AGTGATCTTC 660
55 CAGAAGGTTA AAGAGTGGTC TCTCATGATC ATGAAGACTT TGATAGTTGC GGTGCTGTTG 720
GCTGGAGTTG TCCCTCTCCT TCTGGGCTC CTGTTTGAGC TGGTCATGT GGCTCCCCTG 780
60 AGGGITCCCT TGGATCAGAC TCCTCTTTT TATCCATGGC AGGACTGGGC ACTTGGAGTC 840

CTGCATGCCA AAATCATTGC AGCTATAACA TTGATGGGTC CTCAGTGGTG GTTGAAAAC 900
GTAATTGAAC AGGTTTACGC AAATGGCATC CGGAACATTG ACCTTCACTA TATTGTTTGGT 960
5 AAAGTGGCAG CTCCTGTGAT CTCTGTGCTG TTGCTTTCCC TGTGTGTACC TTATGTCATA 1020
GCTTCTGGTG TTGTTCTTTT ACTAGGTGTT ACTGCGGAAA TGCAAAACTT AGTCCATCGG 1080
10 CGGATTTATC CATTTTACTT GATGGTCGTG GTATTGATGG CAATTTTGTG CTTCCAAGTC 1140
CGCCAGTTTA AGCGCCTTTA TGAACATAAT AAAAATGACA AGTACCTTGT GGGTCAACGA 1200
CTCGTGAAC ACGAACGGAA ATCTGGCAA CAAGGCTCAT CTCCACCACC TCCACAGTCA 1260
15 TCCCAAGAAT AAAGTAGTGT TCTCAACAAC TTGACCTTCC CCTTTACATG TCCTTTTGTG 1320
TGGACTTCTC TCTTTGGAGA TTTTCCAG TGATCTCTCA GCGTTGTTT TAAGTTAAAT 1380
20 GTATTTGACT TGTGTTCTCA GCATTCAGAG AGCAGCGGTG TAAGATTCTG CTGTTCTCCC 1440
TGGATCTTCT GACATTACTG CTGCTGAGA TTTGTATATG GTAAATACA AGTTCTTGA 1500
TACCTTAAA CCTTGGATTA AACAGAATGT GCATTGTACA TCTTTAAACA AAATGTATAT 1560
25 TAATTTATTA AATCTAGTTG TCACTTTAAA AAAAAAAAAA AAAAAACTCG AGGGGGCCCC 1620
GGTACCCAAA T 1631

30

(2) INFORMATION FOR SEQ ID NO: 98:

- 35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 504 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
40 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98:

CCGAGCTGGG CGAGAAGTAG GGGAGGGCAC GAGCGCCGC GGTGGCGGT GCTATCGCTT 60
45 CGCAGAACCT ACTCAGGCAG CCAGCTGAGA AGAGTTGAGG GAAAGTCTG CTGCTGGGTC 120
TGCAGACGCG ATGGATAACG TGCAGCCGAA AATAAACAT CGCCCTTCT GCTTCAGTGT 180
GAAAGGCCAC GTGAAGATGC TCGGCTGGA TATTATCAAC TCACTGGTAA CAACAGTATT 240
50 CATGCTCATC GTATCTGTGT TGGCACTGAT ACCAGAAACC ACAACATTGA CAGTTGGTGG 300
AGGGGTGTTT GCACTTGTGA CAGCAGTATG CTGCTTGCC GACGGGGCCC TTATTTACCG 360
55 GAAGCTTCTG TTCAATCCCA GCGGTCCTTA CCAGAAAAAG CCTGTGCATG AAAAAAAGA 420
AGTTTGTAA TTTTATATTA CTTTGTAGTT TGATACTAAG TATTAAACAT ATTTCTGTAT 480
TCTTCCAAA AAAAAAAAAA AAAA 504
60

(2) INFORMATION FOR SEQ ID NO: 99:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1416 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99:

GGCACGAGGG AGGGAGCCCT CTCGTTGGG TGACTCTTGT GTGCCCTTTA GACAGGCTGG 60
CCTGCCGGTT CCACAGGTA CAGTTAGGAC TTGAGTCTTT CTTTTTCTGT TTTGAGTTGG 120
TGAGTGAGTG ATAGGGTAAC ATGGGCCTTC AGGATGACCC CTTGGAAC TGCCGAGTTC 180
CTTAAATCTC AGCTGGGATC CTGGACCTGG GAGGCCCTG TGAGGGCCAG CTCGTGAAAA 240
ACCTGGGAGT TGATGCCGA GCTGTGAAG AACTCTGCTC GAGGGCAGGG TGCCCTGGAA 300
CACTGGTAGT TCTGGGCTG GGAGGGAGAG GGGCTCCGGC TTTCTCTGAA ATGAACACTG 360
CTCTTCAGCA GTTCAAGTAC TTGTTCTCAA AACATTTTCT AATTGATTGG TAGGTTTTC 420
TAAGCATTTG TTTCTTAAGG CATGGAAGG GAAGAATGCT CAAGCAAGTC ATGTTTGTMT 480
TCAGTGGGAT GGGCCCGCGT TCTCACTGCT GGGGGCTTCC CCTTCATGTG GCACCTTTGT 540
GCAGGGGCCA CCAGGCAGAC TCTTCCCACC TTCTCCCACT GAAGCACCAA GGGGCTTGGA 600
ACCGTAATTT GGCTAATCAG AGGCAITTTT TTTGTCTAG TATCTTTCAC ACTTGTCCAA 660
CCGTCTTATT TTTTAAAG TTCTGTGCT TGTATTAA CAAGAACTAGA GAGAAATAGT 720
TTCTGAAGCC AGTTTATTGT GAAGATCCCC AAGGGGAGGT TCGGTAGAGA AAAATAGTAA 780
GCTGGTTTAG AACTGACGA GGGCAAACAG CCAGGACGCA TTGGAGAGGA ATTTGCCAAA 840
GATCTACCCT GAGATAACGC CTGTCCAGTG TCTTCACCAC GTGAATAACC AGCGCTCCAA 900
AGTGTMTTTC TGCTTTGAAA AAAAAAATC CACAAGCTTT TAAAGGTGCA TTTAAGAATC 960
CATGTGACTT TAGAATGGAA CTGCCGGCCC TGGCAACTGT CAGTGTGCT AGAAGGTTTC 1020
ATGCCTCTGG AATGCATGTG ATACTCATCT CCATTTTGT TCTTGATTG CATTTTGT 1080
CTTTTAGCAG ATCTGTCCCT GTGGGTGGTG TCTAAGAAGT CGGACACCTT GGTTTTGTG 1140
TTAGATTGAG CTGGGAGCT GCAATCAGCT TCTTTATATG CAAATTAGGC ACGACCCATC 1200
TGTGGTTCCT GGTGGTGGC TAATGAAGTG AGGGGAGGGA GGGATGTCAC CCCAAAAGTA 1260
GGCCCTCCCA TTGGCTTTGG CCAGGCCAGA CACTTCACAT CGTTTACATG GTTCTGTGTA 1320
ATTTTAAAGT TTATGTGTAT AAAGCGAAGC TGTTCCTGTG AAAGTGTATA TTTGTAAAT 1380
AAATATATTG CTACTTGAAA AAAAAAAAAA AAAAAA 1416

5 (2) INFORMATION FOR SEQ ID NO: 100:

(i) SEQUENCE CHARACTERISTICS:

10 (A) LENGTH: 2847 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 100:

15 GGCTAGGACA ATTTTGGTGC TTACCTATC TCTGCAAAGA CTGGAGAATT TGGCATACCA 60
TTAATTACAA CCACCAATCA TATCCAACAA AAGTACCCTA AAAGAAGGAC CAGTGGCCAC 120
TCTCGAAAAA ATTTAAGTAT CAGAAGATTA AAAAGATTTT AGGATTTTGA AGCTTGTATT 180
20 GTCTTTCCCC AATAATCATT GTTTGATCTC CAAATAGTAG CCTTATATTA GCAATRGACA 240
GATCATTTGGT TCTCCATATC TGATCATATG TTACTACTTT GGAATCAGTA TTTGGGCAAA 300
25 TTCAAGCATT TATGCAGTGG ATATAAATGG AAATATAAAA ATATTTGCCA ACCTGTCTCA 360
GTAACCTATC ATATCTCTGT GNATCCTCAA GGAAAGCACT TTTGCTTTTA CTTAGAAAGC 420
30 GTTTCAGATT TGCTTTATAG ACTCCTGCTG TCTTCAGTAC CTGATAAAAC TTTAACCAGG 480
GAAGCAITTA ACACAGTGCA GCAGCTTTTG CCCAGGCTTC TAAGTTCTG CCGGCAGCAT 540
TTATCAATGT AAGAACTAGG ATGCTTCCTG CAGTGGCACT ACCTTCCCCT AGAGCTGGAG 600
35 CATGCTGCTT GGCCTTAAGC CCCAGCATGA TGAGGCTTCC CTCCTGCCAG GTCAGTAAAA 660
GTTAGAGAGC TCAGAAATGG GTCTTGCTG GGTGCAGGTG GCAGGGTTTG CTGAAACCCC 720
TAAAGAGAAG TCACCAAGGG AGGCAGGTAA TGAATGTTTC CAGAATCAGT CKGATACTCA 780
40 TAGCAATTTC TGGCTATCTT TCAAATGTTG AATTTCCTGA TGCTGAGAGG GACTTTGATT 840
TGATATCATT AAATCCAGGA CAGTCCCAAG AAGTGCTTGG AGTCTCGGCT CTGACAGCCC 900
45 AAGAAGGGAA ATAACCTGTA TTAAGGAACA ACTATGAGCC AGGCCCTGAG CTGTCTCTTA 960
GATAATAAAA CAGATGGGGA GTGGAAGAGT CATTTGCTTC AAGTTATACA GCTAGGAAAT 1020
50 ACTCAAGCCA AATCTTGAAC GCAGCTCCCC CTAATTCTGT GGACAGGCAC TTTGTACCAC 1080
ACACCATGGT CCACCTAAAA ACAGAAGGAT AAAAAGACTT CAGGTTTTCC CACTGTGTGC 1140
TGACCATCCC AATTATGAA TCTTCTTCAA AATGACATTT CACAGTTATA GTTAGGGCTC 1200
55 AGAAATGGCA TTGAGGTAGC CTTATTTCTC CCCTTTAGCA GATGCTTTAA GTACACATTG 1260
CTGACTTGAG CCCACCCCA GGAGTTAGGA GAACATTTCC TTTTTCATGC CATCTTCCAT 1320
60 AAATAAGGTG TTTCTTGCC TTCAAAGATA TAGAACTTTG CAGCAGTAGT AAAAGTGAAG 1380

GGTGTTCGTC TCTCTACTCA ACTTTATTTG AAAATGTCTG CAGCTTCACT CCTGTAGAAA 1440
AGGAAATCTT CATATTTTAG TAAACTTAGC CGCCAGTGTA CTCTGTGAGG ATGTGGCAAT 1500
5 TCAAAGTCCA GTGAATCTGG CTCTCTTACT GATTCTGGT TTTAGTGTGT GTGTCGGGG 1560
AGTGTGTACC TATATATAAA GGACAAGTGT GATATGTGTG TATATGTATA TACATACATA 1620
CATGTCCACA CACACACACA CAATATTTGA GAGCTAAGGA AAACCTCAAAG CAGCCCCCTC 1680
10 ATTATCTTGC GTACTACTTC AAAGATTTCT GTCAGCCCTA ATTACAAGTG TCACCATATA 1740
GTTGGGGCTT AGGTACTTGC TTACAGGAAG AGCAATTCCC TAGCAAAGGT CATTAGCTCC 1800
15 TAAGGCACTG AGTCAAAGTG ACAGCCCTGA AGGAAATTGC ACTCCAGCCC TCCTCCAGGA 1860
TGTCTAATAA GATGGGAAAC TTGGATGCCC AGCCATTTTG GTGACCTGAG AGTCTAACTA 1920
CTCCAGTTAG ACCTAAGGGC ACAAATGCAG AATTCATGAC CTGTAGTTG TGGCAGGGTC 1980
20 TAGGAAGTCC TCTCTCCCCA AGTAGAAAAT ATTCTCTTGC CATTCTGAA ATTCCACATT 2040
CATATAATGG CTGTGCAATA CATGCTTCTC AATAAGAAAA TTAACGTCAT GTTTACTGTG 2100
25 TGCTGATCAC ATCAGATTTT TATGTTTAAA AAAATCTCAT TATGGNTTGA GTCCAGCCCA 2160
GCTCTAAGAG AAAAAGAAGG CCCATATGGG AGACTTCAGT CTCATTATTA TTGCCTTTAT 2220
CCAGCAGTGC TTATRAAGCC CCCTACCTG TCCCATTCCA GAAACCATAA GACTCAGGCA 2280
30 GTTCTTGATT CTGGAGGCCT GCCTGGTAAG ATAAGATAGT ATAATTTGGA ACTGAGAACA 2340
TACCAGAAAC AGCAGAACGA GGGCCAGAGC AGAAAAATGA AAATAAGTGG AGACACTTAT 2400
35 GGATACATTG GTGCAAAAAA AGCCACGGGS CCCATACTGG GCTTGATATG ACTTTGAGGG 2460
GACAGCAGAT TAATACTTAA TGAGGGTTAA ACCTGACCAG TCTTTCTACA GTGACAGGCC 2520
ACACTGCATG AATGGGGAGA ACCAATGAAT CCATTGTCCT CTGCCTATTT TCCTGTGCAC 2580
40 AGTCACATTG CCTCCTTAGG AATCTTCCCC TTCCACCCTT TACATTAAAC AAGGGAACAC 2640
TGAATCTTTC AAGGGAATTA CACGTTTGGG TTAATGTTTC AGTATATCAT TTTCACTAG 2700
45 TAAATTATTT TGTAAAGAGAG ATTACTGCT ATCCCAGGAT GTTCGGACTT GGTGCCCTG 2760
TGCATTGGA AATCAATAAA CTATTACTGG AAATGCCAAA AAAAAAAAAA AAAAAAAAAAN 2820
50 NAAAAAATC GAGGGGGGCC CGTACCC 2847

(2) INFORMATION FOR SEQ ID NO: 101:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1394 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 101:

5 GAGATTGGTG GAGGAGAGTA AATAATCTAG AGGCAAGAGT TCAGTGAGGG CCAAGGGGGA 60
CCCCCAGAAA AAGGTATGGA GCTAACTCAT CTCTTTTACA AGGGGTGGCC ATGACTTACT 120
GTTGCAAGT ACTCAGTGTA TATTTAATGT TGATTGTTGA ATTTTAGTTA CGAGAGGGAA 180
10 GAACAATTTT ACTTCTGTCC TTATTTCACT TGCTGAAAAG CTGTGGGACA AAATGTATGG 240
AATAGACAAG GCCACTTTCT TTGTGATTTC TGCTTTTCAT GCATATTATT TTATTTACCC 300
ATAATTTCCA AGAGGTTTGG CGTTCGCTC TCCTGCTTTT TTCTTTTCATC CACCCCTTTC 360
15 CTTTTITTTG AAGGGGGTTA TATATGAGAG TTCATTGAAG AAGTCCAGTG AGGCTGAAGT 420
AAAGGGCAA GATAGGCAG TTAACATAAG AGCACTTTAT TTCTTTGAAG CCTTTCTAAG 480
20 AAAGAAATGG GGTGCGAGT GGCTTGAATC TCCCATGATG TTGGAGGGCA CTTAGTGGGG 540
TTGAAGTATG ACATAATATT TCCCATGGG GAAAGGAGAA TTTCTCTTAG AGGGTGGCAA 600
AATGCCTTTG CCCAGTGTCC CTATTTTAGG CATCTTTTCC TTCCTTATTC CTTCCAGTCA 660
25 GGGTGTGTCC TATACAAAAC TTCCCATCAG TTCTCCTCAA TATCCCCAT TTGTAAATGA 720
TCACTTCTCT TTTCTAAACC CTTTCTCTGT TCAGATCCAT ACAGGATTTG CAAGGGTAGG 780
30 ATCATACATG CAAATGCCCC TTGTTTATCT GTGTCTTCTG CAACTAGTC TCATGAAGAA 840
TTCTGGCGTG CAGCAGGGTA GCTGAAGTTT GGGTCTGGGA CTGGAGATTG GCCATTAGGC 900
NTCNCAGAGA TTCCAGCTCC CTTCCACCAA GCCCAGTCTT GCTACGTGGC ACAGGGCAAA 960
35 CCTGACTCCC TTGGGCCTC AGTTTCCCCT CCCCTTCATG AAATGAAAAG AATACTACTT 1020
TTTCTTGTG GTCTAGCATT GCTGGACACA AAGTGTAGTC ATTATTGTTG TATTTGGTGA 1080
40 TGTGTGAAA ACTGCAGAAG CTCCTGCCT ATAAGAGGAA ATAAGAGAGA AAGTGGAGGA 1140
GAGGGACAAA AGGAGTAATT ATTTGGTATA GATCCACCA TCCCAACCTT TCTCTCCTCA 1200
GTCCCTGCTC CTCATGTTTC TGGTTTGGTG AGTCCTTTGT GCCACCACCC ATAATGCTTT 1260
45 GCATTGCTGC ATCCTGGGAA GGGGTATAT GGTCTCACA GTTGTGTCA TTGTTTITTT 1320
GCATGCTTTC TTAATAAAAA AAAAAAAAAA ATGTTTANAG TTTTATCTTA AAAAAAAAAA 1380
50 AAAAAAAAAA ACCC 1394

55 (2) INFORMATION FOR SEQ ID NO: 102:

(i) SEQUENCE CHARACTERISTICS:

- 60 (A) LENGTH: 794 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 102:

5 GGMRCGAGGC GGAGTAAAGG GACTTGAGCG AGCCAGTTGC CGGATTATTC TATTTCCCTT 60
CCCTCTCTCC CGCCCCGTAT CTCTTTTCAC CCTTCTCCCA CCTCGCTCG CGTACCATGG 120
CGGAGCGTCG GCGGCCACTC AGTCCCATTC CATCTCCTCG TCGTCCTTCG GAGCCGAGCC 180
10 GTCCGCGCCC GCGGCGGCGG GGAGCCCAGG AGCCTGCCCC GCCCTGGGGA CGAAGAGCTG 240
CAGCTCTCTCC TGTGCGGTGC ACGATCTGAT TTTCTGGAGA GATGTGAAGA AGACTGGGTT 300
15 TGTCTTTGGA CACGCTGATC ATGCTGCTTT CCTTGGCAGC TTTCAGTGTC ATCARTGTGG 360
GTTTCTTAMC TCATCCTGGC TCTTCTCTCT GTCACCATCA RCTTCAGGAT CTACAAGTCC 420
GTCATCCAAG CTGTWCAGAA RTCAGAARAA GGCCATCCAW TCCAAAGCCT ACCTGGACGT 480
20 AGACATTACT CTGTCTCAG AAGCTTTCCA TAATTACATG AATGCTGCCA TGGTGACAT 540
CAACAGGGCC CTGAAACTCA TTATTCGTCT CTTTCTGGTA GAAGATCTGG TTGACTCCTT 600
25 GAAGCTGGCT GTCTTCATGT GGCTGATGAC CTATGTTGGT GCTGTTTTTA ACGGAATCAC 660
CCTTCTAATT CTGTCTGAAC TGCTCATTIT CAGTGTCCG ATTGTCTATG AGAAGTACAA 720
GACCCAGATT GATCACTATG TTGGCATCGC CCGAGATCAG ACCAAGTCAA TTGTTGAAAA 780
30 GATCCCAAGC AAAA 794

35

(2) INFORMATION FOR SEQ ID NO: 103:

(i) SEQUENCE CHARACTERISTICS:

40

- (A) LENGTH: 1544 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 103:

45

TTTGCTTGCT AGTCTGAACC AAAGAGTTGT TTGGGCATTT GCTGTGTTGG CCATTTCTGG 60
AGCAAGAGGG TCTTCTTCCT CCTTCCCCCA GCCAGCCAGC TGTCTGGGG CCAGGCTTTC 120
50 CTGGGTGGAA AGAAGTATAC CTTTCCCTGG GGCCCTAGGA TAGCAAAGTG AGCCATAGTG 180
GGCCAGGCTG CCTTCATGC TGGGCCCCAG CCCAGGTCTG CACTCGCCTG GATCACCTTC 240
TTTGAGCCTT AGCCATCTCC TGTCAAGTAG GAATGAACCT GCCAGCCTTC AGGYTCGTTT 300
55 AGCTATGACC ATCTGTGCGG TCAGGTACA CTCAGCTCTC CTCCCCAACT CCAGCAGCCT 360
TTAAGAAGTG TCCCTTTGGC GCCCCCTGGA GGCAGAGCAC TGAGCTGGAC CCTGGGTAGA 420
60 CTCCACAGG GAGGACGGAG CTGGCCTCAG GAGTGGGACA CCCAGACTTG GCAGGGCCTT 480

CAAGAGGCCT GTGTGGGGGC CCCAGGAATC CTTAGCTGAA GCGGGGAGAC TCACTCTCCA 540
TCTCAGGAAA TTCTAGCCCT TGCCCTCAGG GAGCCACGGT TGAGGGTGAG GCCCAACACC 600
5 TGCCTTAGGG CCCTGGGTGG GCAAGTCTGG GCCCTGGGGT AGGGAGGGAG ACTCAGGCCC 660
ACACTTGGGT ATTTTCTAAT TTCAGACAAA CACACACTCA GCGCGCACTC ACTGATTCTT 720
10 ACACATTGCC AAGATTTCAC ACATGTGACC AGGGGCCACC AAAGTCCCTG TGACCTTTGT 780
GACTAGGATC CTAATTTCTC TATTTCTCC TGGGTGCCTG GGTCTGTGTC ACCTGGGGCA 840
GTGTGGATAA TGTTTAGTTC TGTGACACTG TTTTGTGGG GTGGCACCTG GTTCTCCGAT 900
15 GCCTGGGCTG GTGTCAGGCC CAGGACTGTA GTGCTGGGAG CAGTAAAGCT CAGCTCTGTG 960
TAATGAGTGA TGCTATGGCT TGCTCGTGTC TTATGATCCA ATCCTTTTCT ACATCAGCCC 1020
20 TTGTTTGTGTT TTATGGCTAG TCTTATCTGG CCTGGTTATT TCCTTGCGGG GAGGAGAGGG 1080
TTTGCTAATC TGCTCCCAGC CCAACCTATT ACCACCCAC CTCGCTGGGA CCTACTGCTC 1140
GGGAGGCAGC AGACAGGGAG CCACCAGCAG TGGCTTCCTG GCCCTGTGCT GGGGGTGGGG 1200
25 GGAAGCTGGG GGCACATGTG GCCCTTGCTT TCTGAGCAGC TCCCAGTGCC AGGGCTTTGA 1260
GACTTTCCCA CATGATAAAA GAAAAGGGAG GTACAGAAGT TCCAATTCCC TTTTATTATT 1320
30 GCTGGTTGGT ATCTGTAAAT GTTTAATAAA TATCTGAGCA TGTATCTATC AACGCCAAGA 1380
ATTTCAAAGT CTCCTTCAAC AATATGAGGC TTTTAGGATG TTTATATTCC TTCATCCCTC 1440
TTGTTTCCCA GGTTTTCAG GAAAAAAG TCTGGAATTA TAGATACAGC TTATTATTAA 1500
35 ATTTGTTCTT GCATAAAAAA AAAAAAAAAA AACNCNNGG GGGG 1544

40

(2) INFORMATION FOR SEQ ID NO: 104:

(i) SEQUENCE CHARACTERISTICS:

45

- (A) LENGTH: 871 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 104:

50

ACCCACGCGT CCGNCTTGTC CACCCGGGG CGTGGGAGTG AGGTACCAGA TTCAGCCCAT 60
TTGGCCCCGA CGCCTCTGTT CTCGGAATCC GGGTGCTGCG GATTGAGGTC CCGGTTCTTA 120
55 AGGTGGGTCG CTGTCCACCC GGGGCGTGG GAGTGAGGTA CCAGATTCAG CCCATTTGGC 180
CCCGACGCCT CTGTTCTCGG AATCCGGTG CTGCGGATTG AGGTCCCGGT TCCTAACGGA 240
CTGCAAGATG GAGGAAGGCG GGAACCTAGG AGGCCTGATT AAGATGGTCC ATCTACTGGT 300
60

CTGTGCAGGT GCCTGGGGCA TGCAATGTG GGTGACCTTC GTCTCAGGCT TTCCTGCTTT 360
TCCGAAGCCT TCCCCGACAT ACCTTCGGAC TAGTGCAGAG CAAACTCTTC CCCTTCTACT 420
5 TCCACATCTC CATGGGCTGT GCCTTCATCA ACCTCTGCAT CTGGGCTTCA CAGCATGCTT 480
GGGCTCAGCT CACATCTCTG GAGGCCAGCC AGCTTTACCT GCTGTTCTTG AGCCTTACGC 540
TGGCCACTGT CAACGCCCCG TGGCTGGAAC CCCGCACCAC AGCTGCCATG TGGGCCCTGC 600
10 AAACCGTGGG AGAAGGAGCG AGGCCTGGGT GGGGAGGTAC CAGGCAGCCA ACAGGTTCCC 660
GATCCTTAAC GCCAGNTGCG AGAGAAGGAC CCCAAGTACA GTGCTCTCCG CCAGAATTTC 720
15 TTCCGCTACC ATGGGCTGTC CTCTCTTTGC AATCTGGGCT GCGTCCTGAG CAATGGGCTC 780
TGCTCTCGTG GCCTTGCCCT GGAAATAAGG AGCCTCTAGC ATGGGCCCTG CATGCTAATA 840
AATGCTTCTT CAGAAAAAAA AAAAAAAAAA A 871
20

(2) INFORMATION FOR SEQ ID NO: 105:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 404 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

30

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 105:

GGCACGAGTT ATAGCATGGC ATTCATACTT TTGTTTATT GCCTCATGAC TTTTITGAGT 60
35 TTAGAACAAA ACAGTGCAAC CGTAGAGCCT TCTTCCCATG AAATTTTGCA TCTGCTCCAA 120
AACTGCTTTG AGTTACTCAG AACTTCAACC TCCCAATGCA CTGAAGGCAT TCCTTGTCAA 180
40 AGATACCAGA ATGGGTTACA CATTTAACCT GGCAAACATT GAAGAACTCT TAATGTTTTC 240
TTTTTAATAA GAATGACGCC CCACTTTGGG GACTAAAATT GTGCTATTGC CGAGAAGCAG 300
TCTAAAATTT ATTTTTTTAA AAAGAGAAAC TGCCCCATTA TTTTGGTGGG GTTGGTTTIT 360
45 AATTTNTAAT NTGAAAAATT TTTTGGGGT TTTTGGGGCC ATGG 404

50

(2) INFORMATION FOR SEQ ID NO: 106:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1542 base pairs

55

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 106:

60

GTCAGACAGG TGGAGCCGCC GGGGAGGAG TCTCAAAGAG CCAGGCTCCA GGAGAGGAAG 60
GGCTCTRCGA GAGGAGAGAG GAGAGCGCTG GAGAGGAGAG GCTGGAGAGT CCTTAGCCAG 120
5 GATGGAGGCT GTTGTTGAAC TGTACCAAGA GGTGATGAAG CACGCAGATC CCCGGATCCA 180
GGGCTACCCT CTGATGGGGT CCCCTTGCT AATGACCTCC ATTCTCCTGA CCTACGTGTA 240
10 CTTGCTTCTC TCACTTGGGC CTCGCATCAT GGCTAATCGG AAGCCCTTCC AGCTCCGTGG 300
CTTCATGATT GTCTACAAC TCTCACTGGT GGCACCTCTC CTCTACATTG TCTATGAGTT 360
CCTGATGTCG GGCTGGCTGA GCACCTATAC CTGGCGCTGT GACCCTGTGG ACTATTCCAA 420
15 CAGCCCTGAG GCACCTAGGA TGGTTCGGGT GGCCTGGCTC TTCCTCTTCT CCAAGTTCAT 480
TGAGCTGATG GACACAGTGA TCTTTATTCT CCGAAAGAAA GACGGGCAGG TGACCTTCCT 540
ACATGTCTTC CATCACTCTG TGCTTCCCTG GAGCTGGTGG TGGGGGGTAA AGATTGCCCC 600
20 GGGAGGAATG GGCTCTTTC ATGCCATGAT AAACCTCTTC GTGCATGTCA TAATGTACCT 660
GTACTACGGA TTATCTGCCT TTGGCCCTGT GGCACAACCC TACCTTTGGT GGAAAAAGCA 720
25 CATGACAGCC ATTCAGCTGA TCCAGTTTGT CCTGGTCTCA CTGCACATCT CCCAGTACTA 780
CTTTATGTCC AGCTGTAAC ACCAGTACCC AGTCATTATT CACCTCATCT GGATGTATGG 840
CACCATCTTC TTCATGCTGT TCTCCAACCT CTGGTATCAC TCTTATACCA AGGGCAAGCG 900
30 GCTGCCCCGT GCACTTCAGC AAAATGGAGC TCCAGGTATT GCCAAGGTCA AGGCCAACTG 960
AGAAGCATGG CCTAGATAGG CGCCACCTA AGTGCCCTCAG GACTGCACCT TAGGGCAGTG 1020
35 TCCGTCAAGT CCCTCTCCAC CTACACCTGT GACCAAGGCT TATGTGGTCA GGAAGTACCA 1080
GGGGACTGGC CCTCCCTCC CCACAGCTGC TCTACAGGGA CCACGGCTTT GGTTCCTCAC 1140
CCACTTCCCC CGGGCAGCTC CAGGATGTG GCCTCATTGC TGTCTGCCAC TCCAGAGCTG 1200
40 GGGGCTAAAA GGGCTGTACA GTTATTTCCC CCTCCCTGCC TTAAACTTG GGAGAGGAGC 1260
ACTCAGGGCT GGGCCACAA AGGGTCTCGT GGCTTTTTC CTCACACAGA AGAGGTGAGC 1320
45 AATAATGTCA CTGTGAGCC AGTCTCACTC CTCCACCCCA CACACTGAAG CAGTAGCTTC 1380
TGGGCCAAAG GTCAGGGTGG GCGGGGGCT GGAATACAG CCTGTGAGG CTGCTTACTC 1440
AACTTGTC TTAATTAAAA GTGACAGAG AAACANAAA AAAAAAAAAA AAAAACTCGA 1500
50 GGGGGGCCG TACCCAAATC GCCGTATGA TCGTAAACAA TC 1542

55

(2) INFORMATION FOR SEQ ID NO: 107:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2327 base pairs

(B) TYPE: nucleic acid

60

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 107:

5	GGTAGCTCAN TGCAGTGAAA TAGTCTTACT GGAAACAAAG CCCTTTATCA AGAATAATTA	60
	ACTCTTCCT TTTCTTTTG GAGAGGTGCT TTGTTTCTGA TCGGACCAIT TCACTGCAGC	120
10	AAGCAACACA GTATTCTRAG CAGAAGATCG GGAATTGAGG CCATGTTGCG GAGGGCCAGT	180
	RACATTATCT GGACTCTGGA GTGTGAGGAA TATGGACTCC ACTCTTCACT ATATTACAR	240
	CGATTACAGAC TTGAGCAACA ATAGCAGTTT TAGCCCTGAT GAGGAAAGGA GAACTAAAGT	300
15	ACAAGATGTT GTACCTCAGG CGTTGTTAGA TCAGTATTTA TCTATGACTG ACCCTTCTCG	360
	TGCACAGACG GTTGACACTG AAATTGCTAA GCACTGTGCA TATAGCCTCC CTGGTGTGGC	420
20	CTTGACACTC GGAAGACAGA ATTGGCACTG CCTGAGAGAG ACGTATGRGA CTYTGGCCTC	480
	AGACATGCAG TGGAAAGTTC GACGGAATC TAGCATTTCT CATCCACGRG CTTCAGTTA	540
	TTCTTGAGGA TCAATTGACA GCTGCAGATC TGGTTCCAAT TTTTAATGGA TTTTAAAAG	600
25	ACCTCGATGA AGTCAGGATA GGTGTTCTTA AACACTTGCA TGATTTTCTG AAGCTTCTTC	660
	ATATTGACAA AAGAAGAGAA TATCTTTATC AACTTCAGGA GTTTTGGTG ACAGATAATA	720
30	GTAGAAATG GCGGTTTCGA GCTGAACTGG CTGAACAGCT GATTTTACTT CTAGAGTTAT	780
	ATAGTCCAG AGATGTTTAT GACTATTTAC GTCCCATTC TCTGAATCTG TGTGCAGACA	840
	AAGTTTCTTC TGTTCGTGG ATTTCTTACA AGTTGGTCAG CGAGATGGTG AAGAAGCTGC	900
35	ACGCGGCAAC ACCACCAACG TTCGGAGTGG ACCTCATCAA TGAGCTTGTG GAGAACTTTG	960
	GCAGATGTCC CAAGTGGTCT GGTGCGCAAG CCTTTGTCTT TGTCTGCCAG ACTGTCAATG	1020
40	AGGATGACTG CCTTCCCATG GACCAGTTTG CTGTGCATCT CATGCCGCAT CTGCTAACCT	1080
	TAGCAAATGA CAGGTTTCCT AACGTGCGAG TGCTGCTTGC AAAGACATTA AGACAACTC	1140
	TACTAGAAAA AGACTATTTT TTGGCCTCTG CCAGCTGCCA CCAGGAGGCT GTGGAGCAGA	1200
45	CCATCATGCC TCTTCAGATG GACCGTGACA GCGATGTCAA GTATTTTGCA AGCATCCACC	1260
	CTGCCAGTAC CAAAATCTCC GAAGATGCCA TGAGCACAGC GTCTTCAACC TACTAGAAGG	1320
50	CTTGAATCTC GGTGTCTTTC CTGCTTCCAT GAGAGCCGAG GTTCAGTGGG CATTCGCCAC	1380
	GCATGTGACC TGGGATAGCT TTCGGGGGAG GAGAGACCTT CCTCTCCTGC GGACTTCATT	1440
	GCAGGTGCAA GTTGCTTACA CCCAATACCA GGGATTTCAA GAGTCAAGAG AAAGTACAGT	1500
55	AAACACTATT ATCTTATCTT GACTTTAAKG KKWAWKMMWW KCTCAGMSRA TTATAMTTSW	1560
	CWMMRARGSM WYMAAWSCTK SWGCTCYWCC KSRSTGRMK MMRCTCTAGA AYTRGYRGAK	1620
60	CMYYYKSGCT KMWGGAAKS GGCASGAGCC AGAGACCTGC ATTGCTTTCT CCTGGTTTTA	1680

TTTAACAATC GACAAATGAA ATTCTTACAG CCTGAAGGCA GACGTGTGCC CAGATGTGAA 1740
5 AGAGACCTTC AGTATCAGCC CTAACCTCTC TCTCCCAGGA AGGACTTGCT GGGCTCTGTG 1800
GCCAGCTGTC CAGCCCAGCC CTGTGTGTGA ATCGTTTGTG ACGTGTGCAA ATGGGAAAGG 1860
AGGGGTTTTT ACATCTCCTA AAGGACCTGA TGCCAACACA AGTAGGATG ACTTAAACTC 1920
10 TTAAGCGCAG CATATTGCTG TACACATTTA CAGAATGGTT GCTGAGTGTC TGTGTCTGAT 1980
TTTTTCATGC TGGTCATGAC CTGAAGGAAA TTTATTAGAC GTATAATGTA TGTCTGGTGT 2040
TTTTAACTTG ATCATGATCA GCTCTGAGGT GCAACTTCTT CACATACTGT ACATACCTGT 2100
15 GACCACTCTT GGGAGTGCTG CAGTCTTTAA TCATGCTGTT TAACTGTG TGGCACAAGT 2160
TCTCTGTGCC AAATAAAATT TATTAATAAG ATCTATAGAG AGAGATATAT ACACTTTTGA 2220
20 TTGTTTCTA GATGTCTACC AATAAATGCA ATTTGTGACC TGTAAAAAAA AAATAAAAAA 2280
ACTCGAGGGG GGCCCGGTAC CCAAATCGCC GATATGATCT AANCATC 2327

25

(2) INFORMATION FOR SEQ ID NO: 108:

30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1062 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 108:

GGCCGCCGAG GCGCAACAGC CGTCTGTCA GCTCTGGGTC CAACCGGACT AGCGAANATC 60
40 TTCTCTATCC TCATCATCGT CTTCCTCATC CCGATCTCGG TCCAGGTCCC TCTCCCCCCC 120
ACACAAGAGG TGGCGAAGGT CCAGCTGTAG TTCTCTGGA CGTTCTCGAA GATGCTCTTC 180
CTCTCTTCG TCATCATCTT CCTCTCGTC TTCTCATCC TCATCATCCA GTTCTCGAAG 240
45 CCGCTCACGA ATCCCCATCC CCCCGCCGGA GRAAGTGACA GGAGGCGGCG GTACAGCTCT 300
TATCGTTCAC ATGACCATTA CCAAAGGCAA AGAGTGCTAC AAAAGGAGCG TGCAATAGAA 360
50 GAAAGAAGGG TGGTCTTCAT TGGAAAGATA CCTGGCCGCA TGAATCGATC AGAGCTGAAA 420
CAGAGGTTCT CCGTTTPTGG AGAGATTGAG GAGTGCACCA TCCACTTCCG TGTCCAAGGG 480
GACAACTACG GCTTCGTAC TTATCGCTAT GCTGAGGAGG CATTGTCAGC CATTGAGAGT 540
55 GGCCACAAGC TGCGGCAGGC AGATGAGCAG CCCTTTGATC TCTGCTTTGG GGGCCGAAGG 600
SWGTNCTGCA AGAGGAGCTA TTCTGATCTT GACTCCAACC GGAAGACTT TGACCCAGCA 660
CCTGTAAAGA GCAAATTTGA TTCTCTTGAC TTTGACACAT TGTTGAAACA GGCCCAAGG 720
60

AACCTCAGGA GGTAACCTTG GGCCCTTCCC TGCTATCCTT TTTCCTCTTT GGAGGTGCCC 780
 AACCTCCTCC ACCCCCTTCC CCTACTCTAG GGGAGAGAGC TGCTAGTGAG ATGACTGTTT 840
 5 TATAAAGAAA TGGAAAAAAG TGAAATAAAA AATATGTTGA ATCAGATTTT TTAAGAGGG 900
 TATTTGTTTT TTTATAACAG GTATTGAAAC AAGTTAACTT GCATTCTTAT GTAAGATAGG 960
 AGGGCTGAG GGGATCCCCA GTGTTTGGAA CATAAGTCAC TATGCAGACT AATAACATC 1020
 10 AACTAGAGAG NAAAAAATAA AAAAAAATAA ATTTAAAAAA CT 1062

15

(2) INFORMATION FOR SEQ ID NO: 109:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 2539 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 109:

25

GAGAGACTCA CACTTCTTTT CCAATTATCAC TGACGATGTA GTGGACATAG CAGGGAAGA 60
 GCACCTACCT GTGTTGGTGA GGTTTGTTGA TGAATCTCAT AACCTAAGAG AGGAATTAT 120
 30 AGGCTTCCTG CCTTATGAAG CCGATGCAGA AATTTTGGCT GTGAAATTC AACTATGAT 180
 AACTGAGAAG TGGGGATTAA ATATGGAGTA TTGTCGTGGC CAGGCTTACA TTGWCTCTAG 240
 TGGATTTTCT TCCAAAATGA AAGTTGTTGC TTCTAGACTT TYAAGMKMRA TWKCCCMK 300
 35 YWAWCKGAAC AMAMKCTGSW CYTCCWSYGC SKTRRMKRYC GYKSTATRRC WARWKSAYM 360
 CCYGKMTGS RRGTAWYTSK TGCAYKAGGG AACAAATGAG GAAGTTTGTT CTTTTTTCCA 420
 40 TCGATCACCA CAACTGCTTT TAGAACTTGA CAACGTAATT TCTGTCTCTT TTCAGAACAG 480
 TAAAGAAAGG GGTAAAGAAC TGAAGGAAAT CTGCCATTCT CAGTGGACAG GCAGGCATGA 540
 TGCTTTTGAA ATTTTAGTGG AACTCCTGCA AGCACTTGTT TTATGTTTAG ATGGTATAAA 600
 45 TAGTGACACA AATATTAGAT GGAATAACTA TATAGCTGGC CGAGCATTTG TACTCTGAGT 660
 GCAGTGTGAG ATTTTGATTT CATTTGTACT ATTTGTGTTT TAAAAAATGT CCTATCTTTT 720
 50 ACAAGAGCCT TTGGGAAAAA CYCMAGGGG CAAACCTCTG ATGTCTTCTT TGCKKMSRT 780
 ARMTTTTGAY ATRMARYACT RMTKSAYTY AAYGRWGTGA CWSGAWAATA TTAASTYTA 840
 TACAATKAAT YWTRRYTSM KRMAGMYAAT CCGAAAYTGT GGMAAMYAAA CTTGATATTC 900
 55 AAATGAAACT CCCTGGGAAA TTCCGAGAG CTCACCAGGG TAACTTGGAA TCTCAGCTAA 960
 CCTCTGAGAG TTAATAATAA GAAACCCTAA GTGTCCCAAC AGTGGAGCAC ATTATTCAGG 1020
 60 AACTTAAAGA TATATTCTCA GAACAGCACC TCAAAGCTCT TAAATGCTTA TCTCTGGTAC 1080

5 CCTCAGTCAT GGGACAAC TC AAATTCAATA CGTCGGAGGA ACACCATGCT GACATGTATA 1140
GAAGTGACTT ACCCAATCCT GACACGCTGT CAGCTGAGCT TCATTGTTGG AGAATCAAAT 1200
GGAAACACAG GGGGAAAGAT ATAGAGCTTC CGTCCACCAT CTATGAAGCC CTCCACCTGC 1260
CTGACATCAA GTTTTTCCT AATGTGTATG CATTGCTGAA GGTCTGTGT ATCTTCCTG 1320
10 TGATGAAGGT TGAGAATGAG CGGTATGAAA ATGGACGAAA GCGTCTTAAA GCATATTTGA 1380
GGAACACTTT GACAGACCAA AGGTCAAGTA ACTTGGCTTT GCTTAACATA AATTTTGATA 1440
TAAAAACGA CCTGGATTTA ATGGTGGACA CATATATTAA ACTCTATACR AKTAMGTCAG 1500
15 MGCTYYCTAC AKAYRAYTCM SWAWMTGTGG AAARYWSSTA MGMSWGCWKK TAMRRRTMCG 1560
GMWWTYYMK RKTYGAYMYW YGCGWMCAG AAAAAGCCGT AAGGTGTATG TAGACCACTT 1620
20 AATCACTAAA TATCTTGCC TATAGGACTC CATTGAATAC ATTAGCCATT GATAATCTAC 1680
CTGTTTAAAT GGGCCCTGTT TGAACCTCA AGCTTTGAAG ACCTACCTGT TCTTCCAGAA 1740
GAGAACGTTG AAAGTGCCAT GTTTCCTTTT GCGTGATCTC TGTTGATGGC ACTCTGGAAT 1800
25 TGTTCAGT TTAAKTCATT TTAGACATAG CATTATTAT CACTGTGGAT CTCTACTTGT 1860
TGGGTGTAT GAATCTTTG AAGAATATAT TTTGAAGAGG TGTGGGAGGA AGGAATACAT 1920
30 TTTATAAAAT GTTGTAGTGA AGCCACAAT TGACCTTKGA CTAATAGGAG TTTTAAGTAT 1980
GTTAAAAATC TATACTGGAC AGTTACAAGA AATTACCGGA GAAAAGCTTG TGAGCTCACC 2040
AAACAAGGAT TTCAGTGTAG ATTTGTCTT TCTTGAACCT AAAGAAACAA ATGACAAAGT 2100
35 TTGAATGGAA AAGCCTGCTG TTGTCCACA TCTCGTTGCT GTTACATTC CTTTGTGGAG 2160
CCTACATCTT CCTAAGCTTT TTAGCAGGTA TATGTTGAAC ACTTCTGTT CATGGTTGAG 2220
40 ACAGAATCAG AGGCCATGGA TACTGACAAC TGATTGTCT GTTTTTTTC TCTGTCTTT 2280
TCCATGACTC TTATATACTG CCTCATCTG ATTTATAAGC AAAACCTGGA AAACCTACAA 2340
AATAAGTGT GTGGTTTATC TAGAAAAATA TGGAAAATAT TGCTGTTATT TTTGGTGAAG 2400
45 AAAATCAATT TGTATAGTT TATTTCAATC TAAATAAAAT GTGAATTTTG TTWWATTAAA 2460
AATTWGSAC AAABTBHGG GGGDTCCAAA CHTWTCGHG KAAMTCTCT WAARMATYTK 2520
50 ATAAACMSCT TCACAATTC 2539

55 (2) INFORMATION FOR SEQ ID NO: 110:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1751 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

60

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 110:

5	AGCATGAAGC CGATGGCCGT GGTGGCCAGT ACCGTCTGG GCCTGGTGCA AAACATGCGT	60
	GCGTTTGCGG GGATCCTGGT GGTGGTCTAC TACGTATTGG CCATCATTGG GATCAACTTG	120
10	TTTAGAGGCG TCATTGTGGC TCTTCTGGA AACAGCAGCC TGGCCCTGC CAATGGCTCG	180
	GCGCCCTGTG GGAGCTTCGA GCAGCTGGAG TACTGGGCCA ACAACTTCGA TGACTTTGCG	240
	GCTGCCCTGG TCACTCTGTG GAACTTGATG GTGGTGAACA ACTGGCAGGT GTTCTGGAT	300
15	GCATATCGGC GCTACTCAGG CCCGTGGTCC AAGATCTATT TTGTATTGTG GTGGCTGGTG	360
	TCGTCTGTCA TCTGGGTCAA CCTGTTCTG GCCCTGATTC TGGAGAACTT CCTTCACAAG	420
	TGGGACCCCC GCAGCCACCT GCAGCCCCCTT GCTGGGACCC CAGAGGCCAC CTACCAGATG	480
20	ACTGTGGAGC TCCTGTTTCA GGATATTCTG GAGGAGCCCG GGGAGGATGA GCTCACAGAG	540
	AGGCTGAGCC AGCACCCGCA CCTGTGGCTG TGCAGGTGAC GTCCGGGCTG CCATCCCAGC	600
25	AGGGGCGGCA GGAGAGAGAG GCTGGCCTAA CACAGGTGCC CATCATGGAA GAGGCGGCCA	660
	TGCTGTGGCC AGCCAGGCAG GAAGAGACCT TTCCTCTGAC GGACCACTAA GCTGGGGACA	720
	GGAACCAAGT CCTTTGCGTG TGGCCCAACA ACCATCTACA GAACAGCTGC TGGTGCTTCA	780
30	GGGAGGCGCC GTGCCCTCCG CTTTCTTTTA TAGCTGCTTC AGTGAGAATT CCTCTGTCGA	840
	CTCCACAGGG ACCTTTCAGA CAAAAATGCA AGAAGCAGCG GCCTCCCCTG TCCCCTGAG	900
35	CTTCGGTGGT GCCTTTGCTG CCGGCAGCCC TTGGGGACCA CAGGCCTGAC CAGGGCCTGC	960
	ACAGGTTAAC CGTGAGTCTG TCTCATCTAT TCACAGCTGG GAATGATACT AATACCTCCG	1020
	ATTTTAGCCC AGCACCACAG GGTACGTTC AGTTTTCCTC TCTTTCCATA GCTGTAAGGC	1080
40	CCTTCTGGG AATGGTCTC ATTCTCCTTA ATCTATTATT GGGTCAGTTT TCCTGCATGT	1140
	CCCCAGCCTC CCATCACTGC CACCCACTCC CCACAGAGAT GCCCTGCTCA TCCGACTGGG	1200
45	GCTTTGACTC CCACACTGTG TACCCCTCTT GTGTGGACGC CCTGCTGCCA AAACCTTCAG	1260
	CAACAGCTT TCCAAATGGA AGTTGTCACT GTCAGGCCTT TACAATCAGC AACAGCAAAA	1320
	TCTACATGCT GCTGAGGGTC CTGCCTCATT AAGATGCAAT AAATATGTAA GTACATAAAA	1380
50	ACAGCAATAG AAGAAACGTA ATGCTTTATT CTCAAATATG ATGTCTACAT AGAAAAGCCA	1440
	AAATTATTAA GAATAGTAAG AATTACCCCA GCACTTTGGG AGGCCGAGGC GGGTGGATCA	1500
55	TGAGGTCAAG AGATCGAGAC CATCTGGCT AACAGGGTGA AACCCCGTCT CTAATAAAAA	1560
	TACAAAAAAT TGGCCGGGCG CAGTGGCGGG CGCCTGTGGT CCCAGCTACT GGGGAGGCTG	1620
60	AGGCAGGAGA ATGGCGTGAA CCCGGGAAGC GGAGCTTGCA GTGAGCCGAG ATTGCGCCAC	1680

TGCAGTCCGC AGTCCAGCCT GGGCGACAGA GCGAGACTCC GTCTCAAAAA AAAAAAAAAA 1740
AAAAAAAAA A 1751

5

(2) INFORMATION FOR SEQ ID NO: 111:

10

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1117 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:

AATGTTGTGG TGGTAGCATT TGGGTAAATT CTRATTATAG AGTCTCTGG AGAGCAATGT 60
20 CCATAAATA ATCCCAAACA ACATTGTCTT TTTRATGTG TAGTGAACAG CAGAGAATTT 120
CAAAGGACCT TGCTAATATC TGTAAAGACG CAGCTACAGC AGGCATCATT GCCTGGGTGT 180
ATGGGGGAAT ACCAGCTTTT ATTCAATGCTA AACACAATA CATTGAGCAG AGCCAGGCAG 240
25 AAATTTATCA TAACCGGTTT GATGCTGTGC AATCTGCACA TCGTGCTGCC ACACGAGGCT 300
TCATTCTGTTA TGGCTGGCGC TGGGGTGGGA GAACTGCACT GTTGTGACT ATATTCAACA 360
30 CAGTGAACAC TAGTCTGAAT GTATACCGAA ATAAAGATGC CTTAAGCCAT TTTGTAATTG 420
CAGGAGCTGT CACGGGAAGT CTTTTTAGGA TAAACGTAGG CCTGCGTGGC CTGGTGGCTG 480
GTGGCATAAT TGGAGCCTTG CTGGGCACTC CTGTAGGAGG CCTGCTGATG GCATTTTACA 540
35 AGTACTCTGG TGAGACTGTT CAGGAAAGAA AACAGAAGGA TCGAAAGGCA CTCCATGAGC 600
TAAAACTGGA AGAGTGGAAG GGCAGACTAC AAGTTACTGA GCACCTCCCT GAGAAAATTG 660
40 AAAGTAGTTT ACAGGAAGAT GAACCTGAGA ATGATGCTAA GAAAATTGAA GCACTGCTAA 720
ACCTTCCTAG AAACCTTCA GTAATAGATA AACAGACAA GGACTGAAAG TGCTCTGAAC 780
TTGAACTCA CTGGAGAGCT GAAGGGAGCT GCCATGTCCG ATGAATGCCA ACAGACAGGC 840
45 CACTCTTTGG TCAGCCTGCT GACAAATTTA AGTGCTGGTA CCTGTGGTGG CAGTGGCTTG 900
CTCTTGCTTT TTTCTTTTCT TTTTAACTAA GAATGGGGCT GTTGTACTCT CACTTTACTT 960
50 ATCCTTAAAT TTAAATACAT ACTTATGTTT GTATTAACTT ATCAATATAT GCATACATGA 1020
ATATATCCAC CCACCTAGAT TTTAAGCAGT AAATAAAACA TTTGCAAAA GATTAAAGTT 1080
GAATTTTACA GTTAAAAAAA AAAAAAAAAA AAAAAA 1117

55

(2) INFORMATION FOR SEQ ID NO: 112:

60

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1313 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 112:

GGCAGAGGTT TTCCTATATT TTAAGTAAAT TTAAAGTGGC TATCAGAATA TTTATTCTTG 60
TTTGAGACTA CCAACATAAC TACGTGTTGA AGGTGCTTCA CAGAGAATAT ATGCGCTTTA 120
ATGTGAAATA ATTTTCACCA ATGTTGCTAA CTTTAATAAA GTATAAAATT TGTAGAATAT 180
15 TCAGTTAAGT AGTTGGTAAC CCTTTTCTAT TTTAGTAAAA CTTAATGCAT GTTTACTTTT 240
TTTGAAGA TGCAGACAAT CTCCTTGAAC ATGAATTGGG GGCTCTCAAT ATGGCTGCAT 300
TACTACGAAA AGAAGAAAGA GCAAGTCTTC TTAGTAATCT TGGCCCATGT TGTAAAGCGT 360
20 TGTGCTTCAG ACGGGATTCT GCAATTCGAA AGCAGCTTGT TAAAAATGAG AAGGGCACCA 420
TAAACAAGC TTACACGAGT GCTCCAATGG TAGACAATGA ATTACTTCGA TTGAGTCTTC 480
25 GGTATTTTAA GCGGAAGACT ACTTGCCATG CTCCAGGACA TGAAAAGACT GAAGATAATA 540
AATTTTCA GTCCAGTATC CAACAGGAAC TGTGTGTGTC TTAAGACCGA AGTTACAATA 600
TGGTATTTT GGTACTGTCT TCCTTCAGCA GTGCATATC TTTTGCAAAG TTCTTTGGTT 660
30 TGACAAGCAT TAGTGACAAA GGCAGAAAAG ATTTATCAGC CATGCTAAAA GAGTGAAGAA 720
TTTGTATCTT TAGAGACACT AGTTTGGCC AACTTAAGAT TTTACGTTAA TTTTACATA 780
35 GTATTTGACA CTCATGCAAA ATAATGTGAA AACATCTAGA TTTAGTAGTT TATTCTGCGC 840
CTTTTGTTAA AACTGAAGAT TTTGGAAAAT GGTGTCACT GCTCTTCCAG CCTATGAATA 900
TTTGTGTGAA ATGGAACCAT GGATTTATGT CTGGATCATC CATACAGAAC CAACAATTTT 960
40 ATTCAAAAAC AATGTGTCA TCAAAGTAAT TGCTCACATT GTGCAGTACT ATGTTGTACA 1020
GACCACGTGA AAGGAATGC TGGTCTAGCT GCGTGGTAT GTTTATAGGC GAATTTTCAGC 1080
45 AGAAGGAAGC CAAAATAGTT TTTTCCCTTTT GAAAGTTTTT TAAAAATTAT TTCATGGGTC 1140
TTTTTTTAA TTAATATGTG TGCATTGTTA CAATGTATGT TGGGATGTCT TTTGACCCTA 1200
AATGCTTTTT TTGTATCAG AGATTGTGTA CTATTTTAT TTTTAATAAA TGTATCTTCC 1260
50 CTTTTMAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAA 1313

55

(2) INFORMATION FOR SEQ ID NO: 113:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1654 base pairs
(B) TYPE: nucleic acid

60

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 113:

5	ACAGGGACAG AATACTTTCT TTCCTTCCTT CAAGTACAAG AAGGCTTTCT CTACCATTTG	60
	CGTCTACACT TTATTTTAAA AGCTATCCTT TTCTAGTAGT ATTTTATCAT GGCAATGGCA	120
10	TGATGACAAC AACAGTCTTT CATTACAGAC TGAAGGGAAG CATGTCCTTA CTAAAAATAG	180
	TTCTGCTACT TTCCCTCCTA TTATAAGGAA ATTTTACAGA TTCTAAAAAT ACCTTAATTT	240
	TTCTTTGATT TTTATTTTAC CAAGTCACAA ATGCTTTTTT GATGTTTTGA GAATGTGTCT	300
15	CATAGAATCA CAAATACTGA CATTTTATTA GATGATTATT TTCCTAGAAT CCCCAGAG	360
	CAGTGGCAGT CCATGGCTTG GTTGAAGCTA GAAATTTTCC TGCCCCCTGGT GACCTGGTAA	420
20	GCCTCTGCT CGGAACCGTG TGAGTGGGTG AGGAAGATGA GAGATGGTCA GATGGAAGAG	480
	AGRAATACAT GAACTGCTCT GGCCTCTCTG GTTCTGTCTT TGGCCAGAG TTTTGA	540
	GCAGCGGANA TNGACTGACT TCACATGCTC AGCTTTCTCA GCCTTTTGTT TATTTGTG	600
25	TCCTTAGATT TCCTGTGTG AAAAGGGGCA AGAAAAGTAA CTCATCATCT CTAACACACC	660
	ATGGCAGCTT AGCCAGGTAG TCTTAGTGGT GGTGTTTAGG CATAAGATAT GCTGATCATC	720
30	AGTCTCAGGC CACAGTTTCC TTCACTAATC GTCCAGCTTG AGTGTCTGT TCTCTTCTG	780
	CCCATTTCTT TGAACCTCCT GCTCTAGCCT TGGCGGAGGG AGAGTGCTAT TTGCTTTGT	840
	TCTCCCTCTG TCTTAGGAAA AGCCATCTTT AATATAGTTC TTCACCACTG TTGGGGTTGT	900
35	TTTGTGATTT TTTTTCCTT CCGAAGAACT CCTGGTTGTT ATTGGATTTT GTATTTTAAT	960
	ACAAATATTT GAATTTTATA AGCTTGTA CAATATTTAA TTAGTGTGAA AGGAAACAAA	1020
40	GAATGCAGGA AAAATAATTT AATATCAACC TCAGTTGACA AGGTGCTCAG ATTATTCAT	1080
	TCGGGATCCT CCTTTGTGTA GGTTTTGTAG ACAACCCTAG ACCTAAACTG TGTACAGAC	1140
45	TTCTGAATGT TTAGGCAGTG CTAGTAATTT CCTCGTAATG ATTCTGTTAT TACTTTCTTA	1200
	TTCTTTATTC CTCTTCTTC TGAAGATTAA TGAAGTTGAA AATTGAGGTG GATAAATACA	1260
	AAAAGTAGT GTGATAGTAT AAGTATCTAA GTGCAGATGA AAGTGTGTTA TATACATCCA	1320
50	TTCAAAATTA TGCAAGTTAG TAATTACTCA GGGTTAACTA AATTACTTTA ATATGCTGTT	1380
	GAAYCTACTC TGTTCTTGG CTAGAAAAA TTATAACAG GACTTTGTAG TTTGGGAAGC	1440
55	CAAATTGATA ATATTCTATG TTCTAAAAGT TGGGCTATAC ATAAATTATT AAGAAATATG	1500
	GATTTTATTT CCCAGGATAT GGTGTTTATT TTATGATATT ACGCAGGATG ATGTATTGAG	1560
	TAAATCAGT TTTGTAAATA TGTAATATG TCATAAATAA ACAATGCTTT GACTTATTTT	1620
60	CAAAAAAATAA AAAAAATAAA NTTCGAGGGG GGGC	1654

5 (2) INFORMATION FOR SEQ ID NO: 114:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 1171 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 114:

15 GGCAAACTTT CCCCAANGC TTCGAAACTT GCAAGCCGAA ACCTTGAATC GTTAAAAGTT 60
GGGTTCGNC GGC GCCCTGG CCCGAAGAAG CGCAATTGGC GTTCCGCGAA CGTTGGCCCT 120
CAACGGCTCG GCAGCCAGCC ATGTCCTGCA CCCAGGACAG CGGCCCTGGG CTACAAGGAC 180
20 CTGMCCTCA TCTTCCTGCG CCGACCTGCG CGGGTAAGG GGWAGTTTCA GACTGTGAAG 240
GACGTCGTGC TGGACTGCCT GTTGGACTTC TTACCCGAGG GGGTGAACAA AGAGAAGATC 300
25 ACACCACTCA CGCTCAAGGA AGCTTATGTG CAGAAAATGG TTAAAGTGTG CAATGACTCT 360
GACCGATGGA GTCTTATATC CCTGTCAAAC AACAGTGGCA AAAATGTGGA ACTGAAATTT 420
GTGGATTCCC TCCGAGGCA GTTTGAATTC AGTGTAGATT CTTTCAAAT CAAATTAGAC 480
30 TCTCTCTGC TCTTTTATGA ATGTCAGAG AACCAATGA CTGAGACATT TCACCCACA 540
ATAATCGGG AGAGCGCTA TGGCGATTTC CAGGAAGCCT TTGATCACCT TTGTAACAAG 600
35 ATCATTGCCA CCAGGAACCC AGAGGAAATC CGAGGGGGAG GCCTGCTTAA GTACTGCAAC 660
CTCTTGGTGA GGGGCTTTAG GCCCGCCTCT GATGAAATCA AGACCTTCA AAGGTATATG 720
TGTTCCAGGT TTTCATCGA CTCTCAGAC ATTGGAGAGC AGCAGAGAAA ACTGGAGTCC 780
40 TATTTGCAGA ACCACTTGT GGAATTGGA AGACCGCAAG TATGAGTATC TCATGACCCT 840
TCATGGAGTG GTAAATGAGA GCACAGTGTG CCTGATGGGA CATGAAAGAA GACAGACTTT 900
45 AAACCTTATC ACCATGCTGG CTATCCGGGT GTTAGCTGAC CAAAATGTCA TTCCTAATGT 960
GGCTAATGTC ACTTGCTATT ACCAGCCAGC CCCCTATGTA GCAGATGCCA ACTTTAGCAA 1020
TTACTACATT GCACAGGTC AGCCAGTATT CACGTGCCAG CAACAGACCT ACTCCACTTG 1080
50 GCTACCCTGC AATTAAGAAT CATTTAAAAA TGTCCTGTGG GGAAGCCATT TCAGACAAGA 1140
CAGGAGAGAA AAAAAAAAAA AAAAAAAAAA A 1171

55

(2) INFORMATION FOR SEQ ID NO: 115:

60 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 842 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 115:

GGTCTGCGCC	GGAAGTGCAT	GAGCTGCCGA	TGTGGTGCTT	AGTGATTGCG	GTTTCGGTCG	60
10	CTCTCCCCGTG	TTTCCCGGC	TGGGTATTG	CCTCGCACCA	TGGCGCCAA	120
	GGCACGAGAG	GGAAGAAGCA	GATATTGAA	GAGAACAGAG	AGACTCTGAA	180
	CGGATCATAC	TGGGGGCCAA	TGCCATTTAC	TGCCTGTGA	CGTTGGTCTT	240
15	TCTGCCTCAT	TTTGGGCCTG	GTGGCCCTG	GGCTTTAGTC	TGGCAGTGTA	300
	TACCACTCTA	TGAGCTCGAT	GGCACGAGCA	GCGTTCTCTG	AGGATGGGGC	360
20	GGTGGCATGG	ACCTCAACAT	GGAGCAGGGC	ATGGCAGAGC	ACCTTAAGGA	420
	CTGACAGCCA	TCGTGCAGGT	GCTCAGCTGC	TTCTCTCTCT	ATGTCTGGTC	480
	CTGGCTCCAG	GCCGGGCCCT	TTACCTCCTG	TGGGTGAATG	TGCTGGGGCC	540
25	GCAGACAGTG	GCACCCAGC	ACCAGAGCAC	AATGAGAAAC	GGCAGCGCCG	600
	CGGCAGATGA	AGCGGTTATA	GCCATTGACA	TTGTGGCCAC	AGGCCACTGG	660
30	CTCTGTCAGG	GTGCACAGCC	CCTCATGCCT	GGAGCAATGA	GGGTCTAGTC	720
	AAGCAGTCTG	AGGTATTGGG	TATACTTATA	CTCTATAGGG	TCGTTGAATA	780
	AATGTGAAAA	AAAAAAAAAA	AAAAAACTCG	AGGGGGGGCC	GGTACCCAAT	840
35	AT					842

40

(2) INFORMATION FOR SEQ ID NO: 116:

(i) SEQUENCE CHARACTERISTICS:

45

- (A) LENGTH: 1640 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 116:

GGCACGAGGC	GGCGGCAGCG	GTGGCGGCGG	CGCCCCCGG	CGGGAGCCGT	TCCCTTTCCC	60
GTCTGGGAGC	GCGGGGYCGG	GGCCAGGGG	ACCCCGGCC	ACGGAGAGCG	GGAAGAGGAT	120
55	GGATTGCCCG	GCCCTCCCCC	CCGGATGGAA	GAAGGAGGAA	GTGATCCGAA	180
	AAGTGCTGGC	AAGAGCGATG	TCTACTACTT	CAGTCCAAGT	GGTAAGAAGT	240
	GCCTCAGTTG	GCAAGGTACC	TGGGAAATAC	TGTTGATCTC	AGCAGTTTGT	300
60						

TGGAAGATG ATGCCTAGTA AATTACAGAA GAACAAACAG AGACTGCGAA ACGATCCTCT 360
CAATCAAAAT AAGGGTAAAC CAGACTTGAA ATACAACATT GCCAATTAGA CAAACAGCAT 420
5 CAATTTTCAA ACAACCGGTA ACCCAAAGTC ACAAATCATC CTAGTAATAA AGTGAAATCA 480
GACCCACAAC GAATGAATGA ACAGCCACGT CAGCTTTTCT GGGAGAAGAG GCTACAAGGA 540
10 CTTTAGTGCA TCAGATGTAA CAGAACAAAT TATAAAAACC ATGGAACCTAC CCAAAGGTCT 600
TCAAGGAGTT GGTCCAGTAG CAATGATGAG ACCCTTTTAT CTGCTGTTGC CAGTGCTTTG 660
CACACAAGCT CTGCGCCAAT CACAGGGCAA GTCTCCGCTG CTGTGGAAAA GAACCTGCTG 720
15 TTTGGCTTAA CACATCTCAA CCCCTCTGCA AAGCTTTTAT TGTCACAGAT GAAGACTCAG 780
GAAACAGAAG AGCGAGTACA GCAAGTACGC AAGAAATTGG AAGAAGCACT GATGGCAGAC 840
20 ATCTTGTCGC GAGCTGCTGA TACAGAAGAG ATGGATATTG AAATGGACAG TGGAGATGAA 900
GCCTAAGAAT ATGATCAGGT AACTTTGAC CGACTTTCCC CAAGAGAAAA TTCTAGGAA 960
ATTGAACAAA AATGTTTCCA CTGGCTTTTG CCTGTAAGAA AAAAAATGTA CCCGAGCACA 1020
25 TAGAGCTTTT TAATAGCACT AACCAATGCC TTTTATAGAT TATTTTGTAT GTATATACT 1080
ATTATTCAAA AAATCATGTT TATTTTGAGT CCTAGGACTT AAAATTAGTC TTTTGTATA 1140
TCAAGCAGGA CCTAAGATG AAGCTGAGCT TTTGATGCCA GGTGCAATCT ACTGGAAATG 1200
30 TAGCACTTAC GTAAACATT TGTTTCCCC ACAGTTTAA TAAGAACAGA TCAGGAATTC 1260
TAAATAAATT TCCCAGTTAA AGATTATTGT GACTTCACTG TATATAAACA TATTTTATA 1320
35 CTTTATTGAA AGGGGACACC GTACATTCT TCCATCGTCA CTGTAAAGAC AAATAAATGA 1380
TTATATTCCA CAGAAAAAA AAAAAAAW MWSTYGARRR GSRGCMCRSW AYMMARWWCC 1440
40 CCWMRIWRGS MKTCSTMTKA YTTACATTCA ACTCTGATCC CGGGGCTTA GGTTTGACAT 1500
GGGAGGTGGG AGGAAGATAG CGCATATATT TGCAGTATGA ACTATTGCCT CTGGGACGTT 1560
GTGAGGAATT GTGCTTTCAC CAGAATTTCT AAGGATTTCT GGCTTAAATA TCACCTAGCC 1620
45 TGTGGTAATT TTTTTCCT 1640

50 (2) INFORMATION FOR SEQ ID NO: 117:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 952 base pairs
(B) TYPE: nucleic acid
55 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 117:

60 TGAATTTAGN AAACACTTTG GAAACTCAT AACCTCATCA GAACTGCCT TTAGCCACAC

60

GAAGAGTGCA AAGACTTTGT GGACAAAATT GGCCAGTTTC AGAAAATAGT TGGTGGTTTA 540
ATTGAGCTTG TTGATCAACT TGCAAAAGAA GCAGAAAATG AAAAGATGAA GGCCATCGGT 600
5 GCTCGGAAct TGCTCAAATC TATAGCAAAG CAGAGAGAAG CTCAACAGCA GCAACTTCAA 660
GCCCTAATAG CAGAAAAGAA AATGCAGCTA GAAAGGTATC GGGTTGAATA TGAAGCTTTG 720
10 TGTAAAGTAG AAGCAGAACA AAATGAATTT ATTGACCAAT TTATTTTTCa GAAATGAACT 780
GAAAATTTcG CTTTTATAGT AGGAAGGCAa AACAAAAAAa AGCCTCTCAa AACCAAAAAa 840
ACCTCTGTAG CATTCCAGCG GCTTGACCAa TGACCTATGT CACAAGAGGT GCGGTGTAAG 900
15 GAATGCAGCC CCCTGAAGAC AGCACTACAA GTCTGGGGGA GCCAGTTTTA ACATCAGTGC 960
ACAGCTGCTG CTGGTGGCCC TGCAGTGTAa GTTCTCACCT CTTATGCTTA GTTGGAActa 1020
AGCAGTTTGT AAACTTTCAT CCTTTTTTTT GTAAATTCAC AAAGCTTTGG AAGGAGAAGC 1080
20 AATAAATTTT TGTTTTCAAA TGGCTGATG TACCTTTTTT CCTGTGCTC TTGAAATATG 1140
TTTAACTCCT CATGAGAGAA CCCTGGATTC TCTATCCCCT AGTCCACAAa ACAaACCAGG 1200
25 CAGTGGTCAG CAGCTACCTT TNAATTTGGAT CACACACGTG AGTCAGACAG TACCAC 1256

30 (2) INFORMATION FOR SEQ ID NO: 119:

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 1143 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 119:

40 GGCCGTAGCA GCCGGGCTGG TCCTGCTGCG AGCCGGCGGC CCGGAGTGGG GCGGCGGCAT 60
GTACCTTCCA CATTGAGTAT TCAGAAAGAA GTGATCTGAA CTCTGACCAT TCTTTATGGA 120
TACATTAAGT CAAATATAAG AGTCTGACTA CTTGACACAC TGGCTCGAGC AAACATGAAC 180
45 GTTGAGTGTG CCCACAGTGA AGTGAATCCA AATACCGTG TCATGAACAG CCGGGGTATG 240
TGGCTGACAT ATGCATTGGG AGTTGGCTTG CTTCATATTG TCTTACTCAG CATTCCCTTC 300
50 TTCAGTGPTC CTGTTGCTTG GACTTTAACA AATATTATAC ATAATCTGGG GATGTACGTA 360
TTTTTGcATG CAGTGAAAGG AACACCTTTC GAAACTCCTG ACCAGGGTAA AGCAAGGCTC 420
CTAACTCATT GGAACAACCT GGACTATGGA GTACAGTTTA CATCTTCACG GAAGTTTTC 480
55 ACAATTTCTC CAATAATTCT ATATTTTCTG GCAAGTTTCT ATACGAAGTA TGATCCAAC 540
CACTTCATCC TAAACACAGC TTCTCTCCTG AGTGACTAA TTCCCAAAAT GCCACAActa 600
60 CATGGTGTTC GGATCTTTGG AATTAATAAG TATTGAAATG TTTTGAACT GAAAAAAAT 660

TCCTGACCTT CTAGATGAGT AACAAAAAA TGAAATAAGT TCTTGAAAT TAAGCCATTT 120
ATTTTAATTT GCTATTTTTT TCAATGTCT AGGTATCTTT AAATTGTGA TTGTGGAATC 180
5 ATTTTCTGTC CAGATACCTT TATCAAAATT ATTGGCCTCA TGAGAGCTGA AGTAAGTCAG 240
CTTTTGGTG AACTTTAGTG GACTTCTGTG AGATTGTAGT TGTACTTTGT ATCTCTAAAT 300
10 CTAAAGATAG TTTTTTAAAA CTCCCAAAGA AAATCTGCTC TCCTTTCTGA TCTAAAAACT 360
CATCTTTGGG GTAAAGAGTT AAGTGTCOA AGGTGTCTAC AGTTCATGAG GTCAGAGGGA 420
GCTAGCCTGG CACCTGGACT CTGCCCATCC ACAGCTGACA GATTCCAACA GAAGTGTATT 480
15 TAAATTCTCC AGTAGACAAT GCTGGGTAAG GGAGGGGGTA GGGCTGGGTT ATTAAGATAC 540
AGGCTGCTGT ATTTTACATT GGTGTGGGG GAAGGGGAGC CTGGAGAAAA CAAAGTCACT 600
20 ATTCCTTTT TTGAAACAGG AAAAAAATT ATTTTGTGT CAGTAAAAAT GGTAGAGAAT 660
TCCAATGTCC CTAGCCACAA GGGACCAGTT CCACTGAGAA GTGAACAGTG GGAACCAAA 720
ATTTCAGAAA CATTTGGGGA AGGAAAAATT GGCTTTCTCT TAATTGGCAG ATGTTCCAGT 780
25 GGGGSGGGG GGCTCTGTTT TTGTTGGGAT GTGTTATGTT GTATGTACGC ATATATGGAC 840
CGGAGTCTGC TGAGTTTATA AGGTTCCAAA AATATGGTAA AATCTTGGTT TTTGTTAATT 900
30 TATCTCAATA AAAGCCCACT GGRACCTCAA AAAAAAAGA NN 952

35 (2) INFORMATION FOR SEQ ID NO: 118:

(i) SEQUENCE CHARACTERISTICS:

40 (A) LENGTH: 1256 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 118:

45 GACGTCATAG GTAAACAGGC TCTGTATCCG TGGCAGCGGC CGTGGCAGGC TGGCTGGGTA 60
CCGGCTGTCTG CTGACCCAGG AGAAGCTGCC TGTCTACATC AGCCTGGGCT GCAGCGCGCT 120
GCCGCCGCGG GGCCGGCAGC TGAACATGT GCTCTTCAGG GCGGGCACCG TGTTCATTC 180
50 ATCTTTGTAC CCCAGCATC TAGCAGTGT GGCATGTAGT AGGCACTCAA GAAATGTGTG 240
TTGAATGAAC GATGCCTGTG ACAAGCAAGC GGACTTTATT CTTTCCTGAC CCTTGCTCCT 300
55 ATGACACACC TCCTCCTGAC TGCCACTGTC ACTCCTTCAG AGCAGAACTC CTCTAGGGAA 360
CCTGGATGGG AAACAGCCAT GGCCAAGGAC ATCCTGGGTG AAGCAGGGCT AACTTTGAT 420
GAACTGAACA AGCTGAGGGT GTTGACCCA GAGGTTACCC AGCAGACCAT AGAGCTGAAG 480
60

TTTACAGCTA CTGAATTTCT TATAAGGAAG GAGTGGTTAG TAAACTGCAC TGTTTCTSTG 720
ATAATGTGAA ATGAGAAGTA TTTACATTGG AGGGCCAATG GCTGGTCCTT CAAGTGCTGT 780
5 TTTGAAGTGC AGATTTCAT TAAATGATGC CTCTGTTTAA TACACCTGGT ACATTCTGA 840
AGAGGGGCTT TATAAGCAGG CTGGGCAGGC CCAGCTTATA AGTTAAAGGG CATCACAGTG 900
10 AGGGTGTAGT AGATAAATTC AAGGAAATAA GAGATTGTGA AGAACTAGG ACCAGCTTAA 960
CTTATAATGA ATGGGCATTG TGTTAAGAAA AGAACATTTC CAGTCATTCA GCTGTGGTTA 1020
TTTAAAGCAG ACTTACATGT AAACCGGAAT CCTCTCTATA CAAGTTTATT AAAGATTATT 1080
15 TTTATTACCG TAAAAAATAA AAAAAAATAA AAAAAAATAA AAAAAAATAA AAAAAAANA 1140
GAN 1143

20

(2) INFORMATION FOR SEQ ID NO: 120:

25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1782 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 120:

CAGGCCCCGG CCCCCACCC ACGTCTGCGT TGCTGCCCGG CCTGGGCCRG GCCCCAAAGG 60
35 CAAGGACAAA GCAGCTGTCA GGGAACTCC GCCGGAGTCG AATTACGTG CAGCTGCCGG 120
CAACCACAGG TTCCAAGATG GTTTCGGGGG GCTTCGCGTG TTCCAAGAAC TGCTGTGCG 180
40 CCCTCAACCT GCTTACACC TTGGTTAGTC TGCTGCTAAT TGGAAATGCT GCGTGGGGCA 240
TTGGCTTCGG GCTGATTTC AGTCTCCGAG TGGTCGGCGT GGTCAATTGCA GTGGGCATCT 300
TCTGTTCCT GATGCTTTA GTGGTCTGA TTGGAGCTGT AAAACATCAT CAGGTGTTGC 360
45 TATTYTTTA TATGATTATT CTGTACTTG TATTATTGT TCAGTTTCT GTATCTTGG 420
CTTGTTTAGC CCTGAACCAG GAGCAACAGG GTCAGCTTCT GGAGGTTGGT TGAACAATA 480
CGGCAAGTGC TCGAAATGAC ATCCAGAGAA ATCTAACTG CTGTGGGTTT CGAAGTGTTA 540
50 ACCCAAATGA CACCTGTCTG GCTAGCTGTG TTAAAAGTGA CCACTCGTGC TCGCCATGTG 600
CTCCAATCAT AGGAGAATAT GCTGGAGAGG TTTTGAGATT TGTGGTGGC ATTGGCCTGT 660
55 TCTTCAGTTT TACAGAGATC CTGGGTGTTT GGCTGACCTA CAGATACAGG AACCAGAAAG 720
ACCCCCGCGC RAATCCTAGT GCATTCCTTT GATGAGAAA CAAGGAAGAT TTCTTTCTGT 780
ATTATGATCT TGTTCACTTT CTGTAATTTT CTGTTAAGCT CCATTGCCA GTTTAAGGAA 840
60

GGAAACACTA TCTGGAAAAG TACCTTATTG ATAGTGGAAT TATATATTTT TACTCTATGT 900
TTCTCTACAT GTTTTTTTCT TTCCGTTGCT GAAAAATATT TGAAACTTGT GGTCTCTGAA 960
5 GCTCGGTGGC ACCTGGGAAT TTACTGTATT CATTGTGGG CACTGTCCAC TGTGGCCTTT 1020
CTTAGCATTT TTACCTGCAG AAAAAGTTG TATGGTACCA CTGTGTGGT TATATGGTGA 1080
ATCTGAACGT ACATCTCACT GGTATAATTA TATGTAGCAC TGTGCTGTGT AGATAGTTCC 1140
10 TACTGGAAAA AGAGTGGRAA TTTATTAAAA TCAGAAAGTA TGAGATCCTG TTATGTTAAG 1200
GGAAATCCAA ATTCCCAATT TTTTGTGGT TTTTGTAGAA AGATGTGTG TGGTAAAAAG 1260
15 TGTTAGTATA AAAATGATAA TTWACKGTA GTCCTTTATG ATWACACCA TGTATTCTAG 1320
AAATAGTTAT GYCTAGGAA ATTGTGGTTT AATTTTGTGAC TTTTACAGGT AAGTGCAAAG 1380
GAGAAGTGGT TTCATGAAAT GTTCTAATGT ATAATAACAT TTACCTTCAG CCTCCATCAG 1440
20 AATGGAACGA GTTTTGAGTA ATCAGGAAGT ATATCTATAT GATCTTGATA TTGTTTTATA 1500
ATAATTTGAA GTCTAAAAGA CTGCATTTT AAACAAGTTA GTATTAATGC GTTGGCCAC 1560
25 GTAGCAAAAA GATATTTGAT TATCTTAAAA ATTGTAAAT ACCGTTTCA TGAAAGTTCT 1620
CAGTATTGTA ACAGCAACTT GTYAAACCTA AGCATATTTG AATATGATCT CCCATAATTT 1680
GAAATTGAAA TCGTATTGTG TGGCTCTGTA TATTCGTGA AAAAATTAAA GGACAGAAAC 1740
30 CTTTCTTTGT GTATGCATGT TTGAATTAAA AGAAAGTAAT GG 1782

35

(2) INFORMATION FOR SEQ ID NO: 121:

(i) SEQUENCE CHARACTERISTICS:

40

- (A) LENGTH: 610 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 121:

45

GTGGCTGCA GATTGTGGT GCGTCTGAG CCGTCTGTCC TGCGCAAGA TGCTTCAAAG 60
TATTATTAAA AACATATGGA TCCCATGAA GCCCTACTAC ACCAAAGTTT ACCAGGAGAT 120
50 TTGGATAGGA ATGGGGCTGA TGGGCTTCAT CGTTTATAAA ATCCGGGCTG CTGATAAAAG 180
AAGTAAGGCT TTGAAAGCTT CAGCGCTGC TCCTGGTCAT CACAACCAGA TTTACTTGGA 240
GTACATGTA AAGAAAACGT CAGTCTGCCT GTAAATTTCA GCAAGCGTG TTAGATGGG 300
55 AGCGTGGAAC GTCACGTAC ACTTGATAA GTACCGTTTA CTTCATGGCA TGAATAAATG 360
GATCTGTGAG ATGCACTGCT ACCTGGTACT GCTTTCAGTG TGTTCCTCCCT CAGCCCTCCG 420
60 GCGTGTCAAG CATACTCTGA GTAGATAATT TGTCATGCAG CGCATGCAAT CAGAATCTCA 480

CTGAGCCACC CATCATTTGTG AAATAATTAC CTCAGTTGTA CAGGACTTGG TGATCAGGAT 540
CCAGGCACTC ACTTGTATTTC TACTGCTCAA TAAACGTTTA TTAAACTTGA AAAAAAAAAA 600
5 AAAAAAAAAA 610

10

(2) INFORMATION FOR SEQ ID NO: 122:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 526 base pairs
15 (B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 122:

GGTACGCCTG CAGGTACCGG TCCGGAATTC CGGGTCGCCC ACGCGTCINGG CCACGCGTCC 60
ACCCACGCGT CCGSCCAGCG GTCGGAGCCG AGCCGGACTG GTCAGGATGA TCACGGACGT 120
25 GCAGCTCGCC ATCTTCGCCA ACATGCTGGG CGTGTGCTC TTCTTGCTTG TCGTTCTCTA
TCACTACGTG GCCGTCAACA ATCCCAAGAA GCAGGAATGA AAGTGGCGCT TTCTCCGCCC 240
CAGGGTTCCA GGACATAGTC TGAGGCAAGA TGGAGGGTAT GAGGGGCCTT CACACTTCAC 300
30 TTCATCCCTT CTACCCATCA CAACATACAA AGCAACTACA CCTGGATTTT TCCAAACAAC
TTTATTTCC TCAGAGTCTT CCTTAATCCT ATGGAACAAG AAGCTGCCAC TGAATAGGGC 420
35 CCAGTATAGG GGCTTGCITT TCTACTCCCT CCCCCAATA TAAAAATATA GACTTTTTAA 480
AAAAAAAAA AAAAAATTCG NGGGGGGSCC GGTACCCATC CCCCTA 526

40

(2) INFORMATION FOR SEQ ID NO: 123:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2081 base pairs
45 (B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 123:

TGTACCGGTC CGGAAATTC CGGGTCGACC CACGTCGTCS GGGGAACATG GCGGCTKCGG 60
AGCCGCGCGT CCTGCGCTC CCCAACAGCG GCGCCGGGGG CGCGGGGGCG CCGTCGGGCA 120
55 CAGTCCCGGT GCTCTTCTGT TTCTCAGTCT TCGCGGACC CTCGTCGGTG CCACACGGGG
CGGGCTACGA GCTGCTCATC CAGAAGTTCC TCAGCCTGTA CGGCGACCAG ATCGACATGC 240
60 ACCGCAAATT CGTGGTGCAG CTGTTGCGCG AGGAGTGGGG CCAGTACGTG GACTTGCCCA 300

	AGGGCTTCGC GGTRAGCGAG CGCTGCAAGG TGCGCCTCGT GCCGYTGCAG ATCCAGCTCA	360
	CTACCCCTGGG AAATCTTACA CCTTCAAGCA CTGTGTTTTT CTGCTGTGAT ATGCAGGAAA	420
5	GGTTCAGACC AGCCATCAAG TATTTTGGGG ATATTATTAG CGTGGGACAG AGATTGTTGC	480
	AAGGGGCCCC GATTTTAGGA ATTCTTGTTA TTGTAACAGA ACAATACCCT AAAGTCTTG	540
10	GGAGCACGGT TCAAGAAATT GATTTAACAG GTGTAAACT GGTACTTCCA AAGACCAAGT	600
	TTTCAATGGT ATTACCAGAA GTAGAAGCGG CATTAGCAGA GATTCCCGGA GTCAGGAGTG	660
	TTGTATTATT TGGAGTAGAA ACTCATGTGT GCATCCAACA AACTGCCCTG GAGCTAGTTG	720
15	GCCGAGGAGT CGAGGTCAC ATTGTTGCTG ATGCCACCTC ATCAAGAAGC ATGATGGACA	780
	GGATGTTTGC COTCGAGCGT CTCGCTCRAR CCGGGATCAT AGTGACCACG AGTGAGGCTG	840
20	TTCTGCTTCA GCTGCTAGCT GATAAGGACC ATCCAAAATT CAAGGAAATT CAGAATCTAA	900
	TTAAGCGCAG TGCTCCAGAG TCGGGTCTGC TTTCCAAAGT ATAGGACATT TGAAGAACTG	960
	GTATGCTACT CACTGGTGAA GGACAGTCAG GTGAAGGACT GTAAGCCAC ACAAGCTCTT	1020
25	CTTATCTCTA CTAGAATTAA AATGTTAAGT CAAAAACGGC TCCTTTTTTG CGCCTCCTAG	1080
	TGAAACTTAA CCAGCTAGAC CATTTGAGTA CCAGCATTTA GTTACAAACG TCAAAGGCTT	1140
30	CCGGTGCTGC TTACCTTCCT TTTTTGTTAA TGTGCTTTTA TTTATTAAAA AAAATTACAA	1200
	TGAAGATGCC TGTTTTGTCT CTACTGTGTA CTCTGATCGT ATCTTTCCAA AGTGCAGACT	1260
	CTTGTAAGT TTTCTTAAAT TGTTCACCTT AAAGAAAATG ACGTACCAAC AATGATTGG	1320
35	CTTTTATATT ACTGTAAGAT GTTATAATGT TAATGTGGAT GTAGTGCTTT TACTTTACAG	1380
	ATTGATTGGA ATAAGATTAT TGCATATGAA TTTACCCACA GGACTCTGAA TCATGTTACC	1440
40	CACTCCCCTC ACAATGTTGT CCACCTAGTG AGTTGCATTG ATCTATCCGT ACCAAATGAT	1500
	GTGAATAAT TACATATCTT TCTTGACTAT ACTGATTCTT TATTTTGGTC ACTATTACTA	1560
	AATCTCTGTT AATATCTCTT CTTTTAACTG AAAAGGGATG GGATAGAAGG GTTTGCAATG	1620
45	CCATATTATT GGTGGAGGGC TGTTTTAACA TCTTTGAAGT ATGGCTTGCT GAATATCTTT	1680
	ACCAACATCT TGAATATATA TTCTAGTGTC CACAAGATTT AGCAAAAAGA TAAAGCTTGG	1740
50	GTGGAATATC ATTTTAAAT GTTCATGTTT TGTTCTATAT TTTCTTCACC TACTCTCCAA	1800
	ATATTGTAAT GCAAAAAGTC TCAGTAATGA TTTGGTAGTA TTAATTTTGT GGTCAATTGT	1860
	TCTCTTCGAT AAATTTATTT TCATTAAATA CTTRTTAGAG GGTTTTGAAA TGTTTTCAA	1920
55	ATATGTGAAA TGTGAACTG CTGTCTTTTA TATTAAAGTA ATTAAGAAA ATGTATTGTG	1980
	ATTGAAATTA TTTTGNCTC CACAAGATGG CTCTATGAGT ATTCTTCAG GAATCTAAT	2040
60	ATTTATTTAA GGTNATAAAA TCTTGACATT TATAATCTTT C	2081

5 (2) INFORMATION FOR SEQ ID NO: 124:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 1717 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 124:

15 CCCCCGCGGA GCTGGACCCG CCGTGGGCTA GGGGCAGGGC CGGAGCCGCG GCGGCGGAGC 60
TGTTGGATCCT TCATGATGAG AGATTTGGGG ACACCTTCTCT CTCCTGTGTG TAGTTGATAG 120
TTTGGTGGTG AAGAGATGCC TGACAGTGTG AAAACCTTTC TCCAGGACCT TGCCAGAGGA 180
20 ATCAAAGACT CCATCTGGGG TATTTGTACC ATCTCAAAGC TAGATGCTCG AATCCAGCAA 240
AAGAGAGAGG AGCAGCGTCG AAGAAGGGCA AGTAGTGTCT TGGCACAGAG AAGAGCCCAG 300
25 AGTATAGAGC GGAAGCAAGA GAGTGAGCCA CGTATTGTTA GTAGAAATTT CCAGTGTGTG 360
GCTTGGAAATG GTGGAGTGTG CTGGTTCAGT CTCCTCTTGT TTTATCGAGT ATTTATTCCT 420
GTGCTTCAGT CGGTAACAGC CCGAATTATC GGTGACCCAT CACTACATGG AGATGTTTGG 480
30 TCGTGGCTGG AATTCCTCCT CACGTCAATT TTCAGTGCTC TTTGGGTGCT CCCCTTGTTT 540
GTGCTTAGCA AAGTGGTGAA TGCCATTGGG TTTCAGGATA TAGCTGACCT GGCATTGAG 600
35 GTATCAGGGA GGAAGCCTCA CCCATTCCCT AGTGTGAGCA AAATAATTGC TGACATGCTC 660
TTCAACCTTT TGCTGCAGGC TCTTTTCTC ATTCAAGGAA TGTGTGTGAG TCTCTTTCCC 720
ATCCATCTTG TCGGTGAGCT GGTAGTCTC CTGCATATGT CCCTTCTCTA CTCACTGTAC 780
40 TGCTTTGAAT ATCGTTGGTT CAATAAAGGA ATTGAAATGC ACCAGCGGTT GTCTAACATA 840
GAAAGGAATT GGCCTTACTA CTTTGGGTTT GGTTCGCCCT TGGCTTTTCT CACAGCAATG 900
45 CAGTCTCAT ATATTATCAG TGGCTGCCTT TTCTCTATCC TCTTTCCMTT ATTCATTATC 960
AGCGCCAATG AAGCAAAGAC CCCTGGCAAA GCRTATCTCT TCCAGTTGCG CCTCTTCTCC 1020
TTGGTGGTCT TCTTAAGCAA CAGACTCTTC CACAAGACAG TCTACCTGCA GTCGGCCCTG 1080
50 AGCAGCTCTA CTTCTGCAGA GAAGTTCCTT TCACCGCATC CGTCGCCTGC CAACTGAAG 1140
GCTACTGCAG GTCAGTGTG TGCCTGCCAT CCAAAGGGGA TGGGCGGGAT TGAAGAAGC 1200
55 TGTGGCAGCT CTTTTCCTG TTCACCTCCC GCCTGCCAGG GAAGGCAGGA CCCGCTCTGC 1260
CAAGGGCCCT CTGCGTATTC CCTTCTCTCT GAGGAATTGA AATTTTGTG TCTGGTGAC 1320
GTAAGGCAGA ATGTTCCCTG ACACCAGTGT GTGGATTTT AACATCACCG TGAGTCTGAA 1380
60

AGGACCACAG GTTTTCTGTC AGCTATTTTC TAGCATTTGC CAGTCCCTGT GCCTGGACTG 1440
ATTGGAACAC TTTGTTTTC TCCCTGTGCC ATTTACCCCTT CCACCTTTC ATCCTGCCTT 1500
5 CTACCACCCT TGGATGAATG GATTTGTAA TTCTAGCTGT TGTATTTGT GAATTTGTTA 1560
ATTTTGTGT TTTCTGTGA AACACATACA TTGGATATGG GAGGTAAAGG AGTGTCCAG 1620
TTGCTCCTGG TCACTCCCTT TATAGCCATT ACTGTCTTGT TTCTTGTAAC TCAGGTTAGG 1680
10 TTTTGGTCTC TCTGTCTCA CTGCAAAAAA AAAAAA 1717

15

(2) INFORMATION FOR SEQ ID NO: 125:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 804 base pairs
20 (B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 125:

CCACGCGTCC GGTCACTATG TAGTGGAGGG GCAGACACCC TCCCGCAAAT TCTGGAAGGT 60
TCTTAGTCTC GACTAGGGCA GTAGCCCCAG GACTCCTAGT CGCCGGCTTC AGGTCACTGC 120
30 CGGCTGAACG GAGCTGCCGT CGCCATGTTT GGCTGCTTGG TGGCGGGGAG GCTGGTGCAA 180
ACAGCTGCAC AGCAAGTGGC AGAGGATAAA TTTGTTTTTG ACTTACCTGA TTATGAAAGT 240
ATCAACCATG TTGTGGTTTT TATGCTGGGA ACAATCCCAT TTCCTGAGGG AATGGGAGGA 300
35 TCTGTCTACT TTTCTTATCC TGATTCAAAT GGAATGCCAG TATGGCAACT CCTAGGATTT 360
GTCACGAATG GGAAGCCAAG TGCCATCTTC AAAATTTCAG GTCTTAAATC TGGAGAAGGA 420
40 AGCCAACATC CTTTGGAGC CATGAATATT GTCCGAACTC CATCTGTTC TCAGATTGGA 480
ATTCAGTGG AATTATTAGA CAGTATGGCT CAGCAGACTC CTGTAGGTAA TGCTGCTGTA 540
TCCTCAGTTC ACTCATTCAC TCAGTTCACA CAAAAGATGT TGGACAATTT CTACAATTTT 600
45 GCTTCATCAT TTGCTGTCTC TCAGGCCAG ATGACACCAA GCCCATCTGA AATGTTCAAT 660
CCGGCAAATG TGGTCTGAA ATGGTATGAA AACTTTCAA GACGACTAGC ACAGAACCTT 720
50 NNTTTTGGN AACATAATT TGAATAAAAT AATTTTAAAT GGATTNTGNA AAAAAAAAAA 780
AAAAAAAAA AAAAAAAAAA AAAA 804

55

(2) INFORMATION FOR SEQ ID NO: 126:

(i) SEQUENCE CHARACTERISTICS:
60 (A) LENGTH: 431 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 126:

GGCACAGCCC AGGGCCTTGA AGCCAGCTGG CCCTGGAGAG GGGCTGCTGT GCCAGCTTGG 60
GGAGGGTCTG GGATGGGGCT GCCCCTGATG GCCCTGATGT GGAGTACCTT GCCAGCATCT 120
10 GCTGGGGTGA ACTTTATTTT AGCCCTTCCC TTGTTGCTCT TATGGAAGAA CAGAGGAGGG 180
GTGGGCAGGT CAGTGATGTC AGCAGTGGAG TGATTCCCAG CACAGCGGCT TCTGGGAAGA 240
15 GGGCATGGAG GCATTTCTTT CAGGGAAATG GTCCATNATT TCAGCCAGAA GGCATTGCAT 300
TAAGTTAAGT CCNGGACTTT TGTGGCCAG CTCTGTGTTA TTAAGGGCCC TTGGCGAAGA 360
CTTCAAGGAG GGGGCAAAAN GACCTTTAAG TTTTtaggTt TAACACAGGG AACCCNCAA 420
20 GGGTTATTTT G 431

25

(2) INFORMATION FOR SEQ ID NO: 127:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 3752 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 127:

NGGCACGAGG AGAGTCACCT GGA CTAGAGATAT CCAATGACCC AGACAAAATT 60
AAACTTCAGC TTTCTAAGCA TAAGGAGTTT CAGAAGACTC TTGGTGGCAA GCAGCCTGTG 120
40 TATGATACCA CAATTAGAAC TGGCAGAGCA CTGAAAGAAA AGACTTTGCT TCCCGAAGAT 180
ASTCAGAAAC TTGACAATTT CCTAGGAGAA GTCAGAGACA AATGGGATAC TGTTTGTGGC 240
AAGTCTGTGG AGCGGCAGCA CAAGTTGGAG GAAGCCCTGC TCTTTTCGGG TCAGTTTCATG 300
45 GATGCTTTGC AGGCATTGGT TGA CTGTTGTTA TACAAGGTGG AGCCACAGCT GGCTGAGGAC 360
CAGCCCGTGC ACGGGGGACC TTGACCTCGT CATGAACCTC ATGGATGCAC ACAAGGTTTT 420
50 CCAGAAGGAA CTGGNGAAAG CGAACAGGAA CCGTTCAGGT CCTGAAGCGG TCAGGCCGAG 480
AGCTGATTGA GAATAGTCGA GATGACACCA CTGGGTAAA AGGACAGCTC CAGGAACTGA 540
GCACTCGCTG GGACACTGTC TGTAACTCT CTGTTTCCAA ACAAAGCCGG CTTGAGCAGG 600
55 CCTTAAAACA AGCGGAAGTG TTTCGAGACA CAGTCCACAT GCTGTTGGAG TGGCTTTCTG 660
AAGCAGAGCA AACGCTTCGC TTTCGGGGAG CACTTCCTGG ATGACACAGA GGCCCTGCAG 720
60 TCTCTCATTG ACACCCATAA GGAATTCATG AAGAAAGTAG AAGAAAAGCG AGTGGACGTT 780

	AAGTCAGCAG TAGCCATGGG AGAAGTCATC CTGGCTGTCT GCCACCCCGA TTGCATCACA	840
	ACCATCAAAC ACTGGATCAC CATCATCCGA GCTCGCTTCG AGGAGGTCCT GACATGGGCT	900
5	AAGCAGCACC AGCAGCGTCT TGAAACGGCC TTGTCAGAAC TGGTGGCTAA TGCTGAGCTC	960
	CTGGAAGAAC TTCTGGCATG GATCCAGTGG GCTGAGACCA CCTCATTTCA GCGGGATCAG	1020
10	GAGCCAATCC CGCAGAACAT TGACCGAGTT AAAGCCCTTA TCGCTGAGCA TCAGACATTT	1080
	ATGGAGGAGA TGAATCGCAA ACAGCCTGAC GTGGACCGGG TCACCAAGAC ATACAAAAGG	1140
	AAAAACATAG AGCCTACTCA CGCGCCTTTC ATAGAGAAAT CCCGCAGCGG AGGCAGGAAA	1200
15	TCCCTAAGTC AGCCAACCCC TCCTCCCATG CCAATCCTTT CACAGTCTGA AGCAAAAAAC	1260
	CCACGGATCA ACCAGCTTTC TGCCCGCTGG CAGCAGGTGT GGCTGTTAGC ACTGGAGCGG	1320
20	CAAAGGAAAC TGAATGATGC CTGGATCGG CTGGAGGAGT TGAAAGAAAT TGCCAACTTT	1380
	GACTTTGATG TCTGGAGGAA AAAGTATATG CGTTGGATGA ATCACAAAAA GTCTCGAGTG	1440
	ATGGATTTCT TCCGGCGCAT TGATAAGGAC CAGGATGGGA AGATAACACG TCAGGAGTTT	1500
25	ATCGATGGCA TTTTAGCATC CAAGTTCCCC ACCACCAAGT TAGAGATGAC TGCTGTGGCT	1560
	GACATTTTCG ACCGAGATGG GGATGGTTAC ATTGATTATT ATGAATTTGT GGCTGCTCTT	1620
30	CATCCCAACA AGGATGCGTA TCGACCAACA ACCGATGCAG ATAAAATCGA AGATGAGGTT	1680
	ACAAGACAAG TGGCTCAGTG CAAATGTGCA AAAAGGTTTC AGGTGGAGCA GATCGGAGAG	1740
	AATAAATACC GGTCTTCCT CGGCAATCAG TTTGGGGATT CTCAGCAGTT GCGGCTGGTC	1800
35	CGTATTCTGC GCAACCGTGA TGGTTGCGT TGGTGGAGGA TGGATGGCCT TGGATGAATT	1860
	TTTAGTGAAA AATGATCCCT GCCGAGCAG AGGTAGAACT AACATTGAAC TTAGAGAGAA	1920
40	ATTCATCCTA CCAGAGGGAG CATCCCAGGG AATGACCCCC TTCGCTCAC GGGGTGGAAG	1980
	GTCCAAACCA TCTTCCCGGG CAGCTTCCCC TACTCGTTCC AGCTCCAGTG CTAGTCAGAG	2040
	TAACCACAGC TGTACATCCA TGCCATCTTC TCCAGCCACC CCAGCCAGTG GAACCAAGGT	2100
45	TATCCCATCA TCAGGTAGCA AGTTGAAACG ACCAACACCA ACTTTTCATT CTAGTCGGAC	2160
	ATCCCTTGCT GGTGATACCA GCAATNAGTT CTTCCCCGGC CTCCACAGGT GCCAAAATA	2220
50	ATCGGGCAGA CCTAAAAAG TCTGCCAGTC GCCCTGGGAG TCGGGCTGGG AGTCGAGCCG	2280
	GGAGTCGAGC CAGCAGCCGG CGAGGAAGTG ACGCTTCTGA CTTTGACCTC TTAGAGACGC	2340
	ATTGCTTGTT CCGACACTTC AGAAAGCAGC GCTGCAGGGG GCCAAGGCAA CTCCAGGAGA	2400
55	GGGCTAAACA AACCTTCAA AATCCCAACC ATGTCTAAGA AGACCACCAC TGCTTCCCCC	2460
	AGGACTCCAG GTCCCAAGCG ATAACACTGT CTAAGCACCC CCAAGCCACT ATCCACTTTG	2520
60	AATCCTGCTC CATACATTGG GTGTATATTT ATTCTGAACG GGAGAAGTTA TATTGTTAAA	2580

AGTGTAAG AATAATTGTG TTATGAAGCT GCCTTATTTT TTTCTTTTTT GTAAGTTACT 2640
ATTTTCATGT GAATATTTAT GTAGATAAAA TTGCTCTCTT GGTAACCTTG TAATGGATGG 2700
5 GGGCCAGAAA TGAAATATTT GAGAAAAACA AGTGAAAAGG TCAAGATACA AATGTGTATT 2760
AAAAAAAAA AAGCCTATTA ATAGGGTTTC TCGCGGTGC AGGGTGTAA ACCTGCTTTA 2820
10 TCTTTTAGGA TTATTCCTAA ATGCATCTTC TTTATAAACT TGACTTGCTA TCTCAGCAAG 2880
ATAAATTATA TTAAAAAAT AAGAATCTG CAGTGTTTAA GGAATCTTTT TTTGTAAAT 2940
CACGGACACC TCAATTAGCA AGAAGTGGG GGAGGGCTTT TTCCATTGTT TAATGTTTG 3000
15 TGATTTTAG CTAAGAGAG GGAACCTCAT CTAAGTAACA TTGACATG ATACAGCAAA 3060
AGGAGTTCAT TGCAATACTG TCTTTGGATA TTGTTTCAGT ACTGGGTGTT TAAAGGACAA 3120
20 ATAGCTGCTA GAATTCAGGG GTAAATGTAA GTGTTTCAGAA AACGTCAGAA CATTTGGGGT 3180
TTTAAACTGA TTTGTTGCTC CCTATCCAGC CTAGACACCA GTAATCTTTG TGTTCACCAG 3240
GACCCAGACC CTTGGCAAGG GATAGGCTCG TTGGTGACAT TGTGAATTC AGATTGTTT 3300
25 TATCCACTTT TTTTGCTATT TATTTAAATG GTCGATCAAC TTCCACAAA CTGAGGAATG 3360
AATTCACGA GCCTGTCTG AAAATGTGGA CGTAAGACAA ACACGTGCTC GTCCTTTAAT 3420
30 GGAGTTCACC AGCACAATG TTAACCACTG CTGTTTGCTT TCGTCTTTT TGTGCGTAA 3480
TAAAGTCAAC TGACCAAGT ACCATGAAAA GGGGCTGCTT GGGGCTCCTG TTTTITAGCT 3540
GCTGTTCTTC AGCTCCGACC ATGTTGCTGT GTGATTATCT CAATTGGTTT TAATTGAGGC 3600
35 AGAAACTGAA GCTCTACCA TGAATGTTT AGAAACAAGA CACACTTTTG TATTAAAAT 3660
GCTTGACGTA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAATCGAGG GGGGCCCGGT 3720
40 ACCCAATTCG CCGTATATGA TCGTAAACAA TC 3752

45 (2) INFORMATION FOR SEQ ID NO: 128:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 1144 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 128:

55 TGACCCCTCTG CCTGCCGGGC TCAGTGCTGG ACGCTTCTG TTTTGTGCA GTCGGTCTC 60
GGTAACACCA GCGGCTGTG GTCCACCACT CCATTAGCA GCTCCATTG GTCCAGCAAC 120
CTTAGCAGCG CCTTCCCTTC ACCACTCCAG CAAACACGCT GGCAAGCATC GGCCTCATGG 180
60

	GCACAGAAAA CTCCCCTGCT CCTCACGCTC CCTCCACCTC CAGTCCAGCT GACGACTTGG	240
	GACAGACCTA CAACCCGTGG CGGATATGGA GCCCCACGAT TGAAGAAGA AGCTCGGACC	300
5	CTTGGTCTAA TTCGCACTTT CCTCACGAGA ATTAAATTAA GCAAAAAACA AACAAACATA	360
	GTGGGCCCTC GTCTAGATCA TGATGTGCCA GTTCTGAGA CATCTTTTTA AGGCTCTTAC	420
	TGCAGCTCCC CTCCCCACCC TCCTCTTCTT TGCAAAACAG ACCCAAGCAG GGCAGGCTCA	480
10	GACCACTCGC TTCTTTCAGA TCTTCTTGC AATTATGATA ACATGAGATT TGCTGTTGTG	540
	CTTTTAGAGA AAAGTCTGGA CTCAGCCACA AACTCTAATA AGACCTGTAC ATCTGAGAAC	600
15	CTTCCCGTT ACTGCGTTTT CACCACCTGT CTTCCCATG CTTTATTTAT CTGTATGAAC	660
	ACAGATTTGA CATTACAGCT AAGGAAATAA TTTGAGTTGA TTCAGAAATC CTGGCATGTG	720
	ACAATTTTGT TAAATTACCA AGTTTGTTTT TTAATAATTT CTCAATATTA TGCGCCAAGA	780
20	TCTAATTTTA AAACGTATG AGGACTTTGT GCTGAAAATA GAGTATTTTT TTAAAGTAAG	840
	GCTGTCTTGG TTAAAAAGCA GATTACAGAA ATGTAAGTCA ACTTAAGAAC RGTGAATGAA	900
25	TGTA AAAACA TTCAGTYGAG ACCATATGCA TTTTCTGTGC TGTTTGTACT TGAGGTATGT	960
	AACATTTGTA TACCTGAAC TATTTTAAAG ATGAAC TGAA ATGCACATAG CCAAGTCTTG	1020
	AGATACAAGA TTGAATGTGT ATTTCTTAAA AATACAAC TT TGTGTGTAC TTTGAAATAA	1080
30	ATGATGCTTT TTTCAAAAAA AAAAAAAAAA AAAAAAAAC TCGAGGGGGG GCCCGGTACC	1140
	CAAT	1144

35

(2) INFORMATION FOR SEQ ID NO: 129:

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1830 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 129:

	GCATGCAGAG GAGCACCTG AGCGTGTGCC TGGAGCAGGC GGCCATSTTG GCACGGAGCC	60
50	ACGGGTGCT GCCCAAGTGC ATCATGCAGG CCACGGACAT CATGCGGAAC AGGGCCCAAG	120
	GGTGAGATT CTGGCCAAA ACCTGCGAGT CAAGGACCAG ATGCCCCAGG GTGCTCCGCG	180
	CCTCTACCGC CTCTGCCAGC CGCCGGTGA TGGGGACCTC TGAACACCCA AATGCCCCAC	240
55	GCTGGGCCGC GGCCTCTGGA GCTGGGATTT GGGAGGACAC AGCAGGCAGC GCTGGCCTTC	300
	TCCAGGATG GCCCAANGCT TCCGCARCCG CCCGTTCCGG GACCTGCCCA GCGTCCTCCC	360
60	TGCCTCCTTC CGGGACAAGC CTGGCCACCC TCGCTGTGAT GACGAGCTGG CTGATTGGCC	420

CTGGGCCGGC CCATTCTTCA CACGCCCTGCC AGAAGCTGGA GGGGTGCTGG AGACCCATAG 480
AGCTGATGGG AGCAGCTGGT GCCTGGCCTT CCGCTCCTGC GTCCCCAGAA CCCAAGGGAA 540
5 CGTCATGGAG GCCACATGGG GCCACCCGGC TCCCTCGGGA TGGCTCCGCT GCACTTTTGA 600
AACCCCGGTT TCCTTCAACG TCCACATTCC AGGTGACCAC ACGTGTCTCC TCCTCCTCAT 660
10 CTTAGCTTCC AGGTTCAACC TAACCTGTGA CTAACCTGCT TGGTGGACTT GGAAAAGACT 720
TGGCTCTGTC GGGAAAGGAG AGACGGGGCC TCCATCACGC CTGTTACCAG AGGATCCCCG 780
AGAGCCACAC CAGCTCTGGA CATCACCGCC CCTGGAAGTG GGGCCACCAG CCCTGGGCAC 840
15 GAGATTTGCT CTGACTTTAT TTATATGGCA TGAAATCTCT GGTTTATTTT GGGATTTTIT 900
GTTGTGGTG TTGTCAAAGT TTGTTTTTTC TAAAGTGTG TGATTATATA TTTGACATT 960
20 TACATTTCAA AGAAAGGTAT GTTGTCTAAC AGGGGACCAA CAGAAGGTAG TATTGACAAC 1020
TGTTCCTGCT TCTACTAAAA AAAAAAGAGC ACAAAGAAA AACTAAATTA TTGAAAAATT 1080
AAAAAATGTC ATTGTTTCCT GTTGTGTAAT ATTAGGGTGT TAAGGTGTCG TTTTGAGGTA 1140
25 TCGACTGTGA TTCTTCCCC CACCCTCCAT TCTCCAGCGG TTGGCCGGTG TTAGAACTCG 1200
CTCTCTTTGA GTGACTGGCT ACAAGGCCCT GAGAGGTGGC CAGCCAGGCT TGGAGCTGGA 1260
30 GGGGATGGAG CCCACCTGA GTGCGCGTGT CACACGGGTT AGAGGGTCAC TGGGAAACAC 1320
CGGGCGGTGG CTTCTGTGAT TTATTTTCTT GATGGTAACT TCTCAGAGCA GGGCRATTGG 1380
GACATCACCA GCCAGAGCAC AGGAAGCCAC CCTGCCTGCT GGGGAGGAGG GACCCACACA 1440
35 AGCCCCCTCG GCAGTTTGTG CCCCCAGCTT CGGTATGCCT TCAGGGAAAG GTCACAGCTG 1500
GGGAGGAAGC GGGGGACGC CTGTCACCCC TGGCAGGTGG TGAGTTCAGG TGGGGGCTCC 1560
40 CTGCTKCCCC CAGGCCTGGG AGCTTGAAGC CCTCCCAGCA TCTGGCATCC GAGCCTCCCG 1620
CCCTCCAGGG TCGCTTCCC TCTCTTGCCG CAGCATACAC GAGGGCAGGC AGTGGCCTTG 1680
TCACTGTATC TTGCATCAGA GACAAAGGAG GACCCGCTTT AGCCCTGCTG CGGGAAATGG 1740
45 GGGATGGCCC AGGGCCAGCG CATTGTGCAC TGGTTTACTT TAAAATGTAC AGATTCTTCT 1800
CGTTAAATTC TTGATAGATT TTTTATTATT 1830
50

(2) INFORMATION FOR SEQ ID NO: 130:

55 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1864 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
60 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 130:

	GGCCGCCCGG ATGGCGACCC CAGCCTCGGC CCCAGACACA CGGGCTCTGG TGGCAGACTT	60
5	TGTAGGTTAT AAGCTGAGGC AGAAGGGTTA TGTCTGTGGA GCTGGCCCCG GGGAGGGCCC	120
	AGCAGCTGAC CCGCTGCACC AAGCCATGCG GGCAGCKGGA GATGAGTTG AGACCCGCTT	180
	CCGGCGCACC TTCTCTGATC TGGCGGCTCA GCTGCATGTG ACCCCAGGCT CAGCCCAACA	240
10	ACGCTTCACC CAGGTCTCCG ATGAACTTT TCAAGGGGGC CCCAACTGGG GCCGCCTTGT	300
	AGCCTTCTTT GTCTTTGGGG CTGCACTGTG TGCTGAGAGT GTCAACAAGG AGATGGAACC	360
15	ACTGGTGGGA CAAGTGCAGG AGTGGATGGT GGCCTACCTG GAGACGCGGC TGGCTGACTG	420
	GATCCACAGC AGTGGGGGCT GGTATCCCA GATCACTGAA GCTGAGATGG CTGATGAAGT	480
	AATTTGCAGT GAAATTTTAA GCGACTGTGA CTCTGCTGCA AGTTCCCCAG ATCTTGAGGA	540
20	GCTGGAAGCT ATCAAAGCTC GAGTCAGGGA GATGGAGGAA GAAGCTGAGA AGCTAAAGGA	600
	GCTACAGAAC GAGGTAGAGA AGCAGATGAA TATGAGTCCA CCTCCAGGCA ATGCTGGCCC	660
25	GGTGATCATG TCCATTGAGG AGAAGATGGA GGCTGATGCC CGTTCCATCT ATGTTGGCAA	720
	TGTGGACTAT GGTGCAACAG CAGAAGAGCT GGAAGCTCAC TTTTCATGGCT GTGGTTCAGT	780
	CAACCGTGTT ACCATACTGT GTGACAAATT TAGTGGCCAT CCCAAAGGGT TTGCGTATAT	840
30	AGAGTTCTCA GACAAAGAGT CAGTGAGGAC TTCCCTGGCC TTAGATGAGT CCCTATTTAG	900
	AGGAAGGCAA ATCAAGGTGA TCCCAAAACG AACCAACAGA CCAGGCATCA GCACAACAGA	960
35	CCGGGGTTTT CCACGAGCCC GCTACCGCGC CCGGACCACC AACTACAACA GCTCCCGCTC	1020
	TCGATTCTAC AGTGGTTTTA ACAGCAGGCC CCGGGGTCGC GTCTACAGG GCCGGGCTAG	1080
	AGCGACATCA TGGTATTCCC CTACTAAAA AAAGTGTGTA TTAGGAGGAG AGAGAGGAAA	1140
40	AAAAGAGGAA AGAAGGAAAA AAAAAAGAAT TAAAAAATA AAAAAAATA ACAGAAGWTG	1200
	MCCTTGATGG AAAAAAATA TTTTATAAAA AAAAGATATA CTGTGGAAG GGGAGAATC	1260
45	CCATAACTAA CTGCTGAGGA GGGACCTGCT TTGGGGAGTA GGGGAAGGCC CAGGGARTGG	1320
	GGCAGGGGGC TGCTTATTCA CTCTGGGGAT TCGCCATGGA CACGTCTCAA CTGCGCAACT	1380
	GCTTGCCCAT GTTTCCTGCG CCCACCCAC CCTCTTCTC CGGCTCCCTG CCCCTCAGA	1440
50	TTGCTGGTG ATCTATTTTG TTTCCTTTTG TGTTCCTTT TCTGTTTTGA GTGTCTTTCT	1500
	TTGCAGGTTT CTGTAGCCGG AAGATCTCCG TTCCGCTCCC AGCGGCTCCA GTGTAAATTC	1560
55	CCCTTCCCC TGGGAAATG CACTACCTTG TTTTGGGGG TTTAGGGTG TTTTGTTTT	1620
	TCAGTTGTTT TGTTTTGTG TTTTNTTNT TTTCCTTTG CTTTTTCCC TTTTATTTG	1680
60	AGGGAATGGG AGGAAGTGGG AACAGGGAGG TGGGAGGTG ATTTTGTTA TTTTTTAGC	1740

TCATTTCCAG GGGTGGGAAT TTTTTTTTAA TATGTGTCAT GAATAAAGTT GTTTTGTGAAA 1800
AKAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1860
5 AAAA 1864

10 (2) INFORMATION FOR SEQ ID NO: 131:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2041 base pairs
(B) TYPE: nucleic acid
15 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 131:

20 GGCACGAGCG CGCGGCAGGG CCTTGGACCC GCGCGGCTCC CGGGGATGGT GAGCAAGGCG 60
CTGCTGCGCC TCGTGTCTGC CGTCAACCGC AGGAGGATGA AGCTGCTGCT GGGCATCGCC 120
TTGCTGGCCT ACGTCGCCTC TGTTTGGGGC AACTTCGTTA ATATGAGGTC TATCCAGGAA 180
25 AATGGTGAAC TAAAAATTGA AAGCAAGATT GAAGAGATGG TTGAACCACT AAGAGAGAAA 240
ATCAGAGATT TAGAAAAAG CTTTACCCAG AAATACCCAC CAGTAAAGTT TTTATCAGAA 300
30 AAGGATCGGA AAAGAATTTT GATAACAGGA GCGCGAGGGT TCGTGGGCTC CCATCTAACT 360
GACAAACTCA TGATGGACGG CCACGAGGTG ACCGTGGTGG ACAATTTCTT CACGGGCAGG 420
AAGAGAAACG TGGAGCACTG GATCGGACAT GAGAACTTCG AGTTGATTAA CCACGACGTG 480
35 TGGAGCCCCT CTACATCGAG GTTGACCAGA TATACCATCT GGCATCTCCA GCCTCCCCCTC 540
CAAACCTACAT GTATAATCCT ATCAAGACAT TAAAGACCAA TACGATTGGG ACATTAAACA 600
40 TGTITGGGGCT GGCAAAACGA GTCGGTGCCC GTCTGCTCCT GGCTCCACA TCGGAGGTGT 660
ATGGAGATCC TGAAGTCCAC CCTCAAAGTG AGGATTACTG GGGCCACGTG AATCCAATAG 720
GACCTCGGGC CTGCTACGAT GAAGGCAAAC GTGTTGCAGA GACCATGTGC TATGCCTACA 780
45 TGAAGCAGGA AGGCGTGAA GTGCGAGTGG CCAGAACTTT CAACACCTTT GGGCCACGCA 840
TGCACATGAA CGATGGGCGA GTAGTCAGCA ACTTCATCCT GCAGGCGCTC CAGGGGGAGC 900
50 CACTCACGGT ATACGGATCC GGTCTCAGA CAAGGGCGTT CCAGTACGTC AGCGATCTAG 960
TGAATGGCCT CGTGGCTCTC ATGAACAGCA ACGTCAGCAG CCCGGTCAAC CTGGGGAACC 1020
CAGAAGAACA CACAATCCTA GAATTTGCTC AGTTAATTAA AAACCTTGTT GGTAGCGGAA 1080
55 GTGAAATTCA GTTCTCTCC GAAGCCCAGG ATGACCCACA GAAAAGAAAA CCAGACATCA 1140
AAAAAGCAAA GCTGATGCTG GGTGGGAGC CCGTGGTCCC GCTGGAGGAA GGTTTAAACA 1200
60 AAGCAATTCA CTACTCCGT AAAGAACTCG AGTACCAGGC AAATAATCAG TACATCCCCA 1260

5 AACCAAAGCC TGCCAGAATA AAGAAAGGAC GGA CTGCGCA CAGCTGAACT CCTCACTTTT 1320
AGGACACAAG ACTACCATG TACACTTGAT GGGATGTATT TTTGGCTTTT TTTTGTGTGTC 1380
GTTTAAAGAA AGACTTTAAC AGGTGTCATG AAGAACAAAC TGGAAATTCA TTCTGAAGCT 1440
TGCTTTAATG AAATGGATGT GCCTAAAAGC TCCCCTCAA AACTGCAGA TTTTGCCTTG 1500
10 CACTTTTGA ATCTCTCTTT TTATGTAAAA TAGCGTAGAT GCATCTCTGC GTATTTTCAA 1560
GTTTTTTTAT CTGCTGTGA GAGCATATGT TGTGACTGTC GTTGACAGTT TTATTTACTG 1620
GTTTCTTTGT GAAGCTGAAA AGGAACATTA AGCGGGACAA AAAATGCCGA TTTTATTTAT 1680
15 AAAAGTGGGT ACTTAATAAA TGAGTCGTTA TACTATGCAT AAAGAAAAAT CCTAGCAGTA 1740
TTGTCAGGTG GTGGTGCGCC GGCATTGATT TTAGGGCAGA TAAAGAATT CTGTGTGAGA 1800
20 GCTTTATGTT TCTCTTTTAA TTCAGAGTTT TTCCAAGGTC TACTTTTGAG TTGCAAACTT 1860
GACTTTGAAA TATTCCTGTT GGTATGATC AAGGATATTT GAAATCACTA CTGTGTTTGT 1920
CTGCGTATCT GGGGCGGGG CAGGTGGGG GGCACAAAGT TAACATATTC TTGGTTAACC 1980
25 ATGGTTAAAT ATGCTATTTT AATAAAATAT TGAAACTCAC CAAAAA AAAA AAAAAA 2040
A 2041

30

(2) INFORMATION FOR SEQ ID NO: 132:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2012 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 132:

TACCAAGCTG CAAGAATCTA CTATATCATG GCAGAAGAAG TAGAGTGGGA CTATTGCCCT 60
45 GACCGGAGCT GGAACGGGA ATGGCACAAC CAGTCTGAGA AGGACAGTTA TGGTTACATT 120
TTCCTGAGCA ACAAGGATGG GCTCCTGGGT TCCAGATACA AGAAAGCTGT ATTCAGGGAA 180
TACACTGATG GTACATTCA GNTCCCTCGG CCAAGGACTG GACCAGAAGA AACTTGGGA 240
50 ATCTTGGGTC CACTTATCAA AGGTGAAGTT GGTGATATCC TGACTGTGGT ATTCAAGAAT 300
AATGCCAGCC GCCCTACTC TGTGCATGCT CATGGAGTGC TAGAATCTAC TACTGTCTGG 360
55 CCACTGGCTG CTGAGCCTGG TGAGGTGGTC ACTTATCAGT GGAACATCCC AGAGAGGTCT 420
GGCCCTGGGC CAATGACTCT GCTTGTGTTT CCTGGATCTA TTATTCGTCA GTGGATCCCA 480
TCAAGGACAT GTATAGTGGC CTGGTGGGC CCTTGGCTAT CTGCCAAAAG GGCATCCTGG 540
60

	NAGCCCCATG GAGGACGGAN TGACATGGAT CGGGAATTG CATTGTTGTT CTTGATTTTT	600
	GATGAAAATA AGTCTTGGTA TTTGGAGGAA AATGTGGCAA CCCATGGGTC CCAGGATCCA	660
5	GGCAGTATTA ACCTACAGGA TGAAACTTTC TTGGAGAGCA ATAAAAATGCA TGCAATCAAT	720
	GGGAAACTCT ATGCCAACCT TAGGGGTCCT ACCATGTACC AAGGAGAACG AGTGGCCTGG	780
10	TACATGCTGG CCATGGGCCA AGATGTGGAT CTACACACCA TCCACTTTCA TGCAGAGAGC	840
	TTCTCTATC GGAATGGCGA GAACTACCGG GCAGATGTGG TGGATCTGTT CCCAGGGACT	900
	TTTGAGGTTG TGGAGATGGT GGCCAGCAAC CCTGGGACAT GGCTGATGCA CTGCCATGTG	960
15	ACTGACCATG TCCATGCTGG CATGGAGACC CTCTTCACTG TTTTCTCTCG AACAGAACAC	1020
	TTAAGCCCTC TCACCGTCAT CACCAAAGAG ACTGAAAAAG CAGTGCCCCC CAGAGACATT	1080
	GAAGAAGGCA ATGTGAAGAT GCTGGGCATG CAGATCCCCA TAAAGAATGT TGAGATGCTG	1140
20	GCCTCTGTTT TGGTTGCCAT TAGTGTCAAC CTTCTGCTCG TTGTTCTGGC TCTTGGTGGG	1200
	GTGGTTTGGT ACCAACATCG ACAGAGAAAG CTACGACGCA ATAGGAGGTC CATCCTGGAT	1260
25	GACAGCTTCA AGCTTCTGTC TTTCAAACAG TAACATCTGG AGCCTGGAGA TATCCTCAGG	1320
	AAGCACATCT GTAGTCACT CCCAGCAGGC CATGGACTAG TCACTAACCC CACACTCAAA	1380
30	GGGGCATGGG TGGTGGAGAA GCAGAAGGAG CAATCAAGCT TATCTGGATA TTCTTTCTTT	1440
	TATTTATTTT ACATGGAAAT AATATGATTT CACTTTTCTT TTAGTTTCTT TGCTCTACGT	1500
	GGGCACCTGG CACTAAGGGA GTACCTTATT ATCCTACATC GCAAATTTCA ACAGCTACAT	1560
35	TATATTTCTT TCTGACACTT GGAAGGTATT GAAATTTCTA GAAATGTATC CTCTCACAA	1620
	AGTAGAGACC AAGAGAAAAA CTCATTGATT GGGTTTCTAC TTCTTTCAAG GACTCAGGAA	1680
40	ATTTCACTTT GAACTGAGGC CAAGTGAGCT GTTAAGATAA CCCACACTTA AACTAAAGGC	1740
	TAAGAATATA GGCTTGATGG GAAATTGAAG GTAGGCTGAG TATTGGGAAT CCAAATGAA	1800
	TTTTGATTCT CCTTGGCAGT GAACTACTTT GAAGAAGTGG TCAATGGGTT GTTGCTGCCA	1860
45	TGAGCATGTA CAACCTCTGG AGCTAGAAGC TCCTCAGGAA AGCCAGTTCT CCAAGTTCTT	1920
	AACCTGTGGC ACTGAAAGGA ATGTTGAGTT ACCTCTTCAT GTTTTAGACA GCAAACCCTA	1980
50	TCCATTAAAG TACTTGTTAG AACACTGAAA AA	2012

(2) INFORMATION FOR SEQ ID NO: 133:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1669 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear



(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 133:

5	GAGCAGTATT TTAACCAACT TGTATTACAG ATGTTACAGT TCATGTTAGG AAGTCAGAAA	60
	AGACTTTGTT TGTCTTTGTT CTGCTGATGT GAGTCATGTT TTGTGGGGTC TTCCATGGCA	120
	CATTTACCTG TTGCTCCGTC CAGATGTTGA GGGCCAGTCT AGGCTGACAC ATCCTACCCG	180
10	AGGACAAGCC TGTCTCCAT TTCTTCACTC TCCCCTCCCC ATATAGCAAC TCTCCAGGT	240
	TTAGATTACC GTTTTCGACG ACAGATTAAAC CAAAAATGCC CCACACAGGT TTTATTACTG	300
	TTATATACTA TACTTTTAAC AGTACAGACC CTAAATTTTA TTATTGTTG CTCCCCAAT	360
15	CTGATACCAA ATGTTTAAAG TTGTTTGAAA TCCAAACATG GTAGTGTTC TGGGTAAATA	420
	TTTTCTAGGC TATGTAAGAG TTAGCAGCCC ATAGCATAGA AGTAATCAAG TAGCATCTGA	480
20	GACTGTGGA GGCAGTAGGG CCTCTCTGGG CCTAACAGCC TCACTTCCCC AGCCTCACCT	540
	TGCTGTCTC TGACACTGCC ATCAGGGCTG TTAGTGGCAC CTGTATGAGG CCAAGTGTG	600
	GTCCAGGGGA ACAGCACAGG TTAATGCGTC TCCCTAGAAC TCATGAAGTC AGTTTAATTC	660
25	ATGCATGAAC ATGAGTTCAT TTTATGTTT ATATAGCTTT CTTAGACATA CCAAACCATC	720
	ATTCATAAAT CAGATAAATT ATTCAGTTTT TGTGTTTGA AAGCTAAGTA TGTGTAGCTG	780
30	GAAACAAAA TGAGCGTGT TTCTCTCTG TTAATCTAGA GTGTGCAGTT ACACATGTGT	840
	GGATAATTT ATGTTCCAGG GGCCTTGGC ATCTCCCAGT GACTGATTCC CAGGAAGAAA	900
35	AGCCCAAAGG GAAACCCAGC ATTCCTTTTCG AGTAGATGTG GGAAAGAGCC CATTGGAGGA	960
	TATGAGGTCC TGTGAAATTC AGTTGTGTGT GTGGCTCCTT GTTAGCAGTC ATGTTGACAT	1020
	GGTGTAGGA GGCTCCCAT CCACCTTTA CATGATGTAG GGACCAAGT CTTGTGAGAT	1080
40	TAACCTTGGG ACACAGTGGG TTAGCCTGGA GAAAATGAGA GGCCCTGCCT GGACCCAGGG	1140
	AGAGGAGCCA GTGACACAGG CAGAGCGGTG CAGCCCTCCT TCCCTTCCAT TTGGAGGAGG	1200
	TGGTGCCAGG AGCCTGCCCC CTTACCTCTG CTGAAGCATA AGTGGACTTT GCTTTTGGGG	1260
45	CTTATCTCTG ATACATGCTG GAGCCCTGCC TCTCCACTGC TAGATGGAAC CTGGAATCTC	1320
	TCATCTACCT CTTAGTCTGT CAGTTTCTAC GTGTGAGAAG CAAGCTTGTG GGCCAGTGTG	1380
50	CTGTGACATG CTGTAGCACT TAAAAATAA TTCCAGGGTT CCCTGGAAAA CCAGTCCAG	1440
	GGTTCCTATG ATCTGTAGTT TCTACCTGGA TTATAACTGG TTTTGGGTAC CTGAATTTTG	1500
	ATTGGTTAGC CTTAATTATA GTCTGGCGTG ATCAIGTAGA ATCTTTTCTG GTGAACAGAT	1560
55	CATAAAGTTC TATCAAGGAG TTCTATCAAG GCATCCATGT CAGTGGTGCT ATGCTGGTTA	1620
	CAACTTGAGA TTTTGAAT AAAAAATTG TCATAAAAAA AAAAAAAA	1669
60		

(2) INFORMATION FOR SEQ ID NO: 134:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1565 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 134:

CACTTTTGCT ATATAACCTA AGTGATAACC CTCCTTTTAGT TACCTGCCAA ACTCTGGNCT 60
15 TGGTTTATAT TGCAGTTAAC ACAGTTACAA AGCTGTAATG GTGTCCTTTT TTCCTTTGTA 120
ACGGAATGTG TAAATCAAAG TATATACATT GTGTGGTGT CCTGTTCTCG GAGTTTCATG 180
20 AGGATTTACA CATGGCATT CAGTGTCTGT ATAGATCTGC CTACCTTTGT GAATTCATCT 240
GTTAACCCCT CTCCTTTGA GAGAGCACCG GCGATGGTGG TTAACCTCTT GTGTTTCTC 300
TCTCTCTAC TGGTTATTCT TGAATTAAGC ACAGACTCGT CAGCTCGGTT GCTTTATCAT 360
25 GAATAATGTG TGTGACCTTG CAGTTCTTCC ACAGTTCAGC AAACAAGTGC TAGCTTCACT 420
GACCAAAAAT TAAGGAAGGA AACACAGTT TTTAAAACGA TCCATCTTTT AACAGCCGAA 480
ACCGATGTGT CTATGGTGCT GCACCTTGCT GTTGTACTTC TGAAATCAGA CGTGTGTGAA 540
30 CGATCATTTT TGACTTAACC GTGAGATGCT CACGAGTACC CTCTCTGTG TTTTGTTAGC 600
ATTGAAATCG AGACTATTTA TTTGGAATAT ATACAACAGT GTTTTTCAC TGTATTTTCA 660
35 TTGCAAAAGT TGAGAACTGC TTTCTCTACC TTTTGCAAAA TAATGATAT TCCATATTGG 720
ATTCTCAAAG ACTTCGATAT GGTGAACCTA TTAAACCTAG AAATGTATT CATCCTTTCA 780
TGACTGTGGC CTGAGTTCCC CAGCCCCCTCT CCTCCTTTT TTTAGATGAG ATTTAGCACA 840
40 CTCTCAGTTA TTTAAACATG CAACATTTCT TGAGTATGTA TGTGAGGCC ATCTGAGCTC 900
ATAGCTGATT CAGTAACCAG TTTTATGCTG TGTCAATCAC ACTCACTACT TAATACTGCC 960
45 ATGGTGAAAA TGTGGAGGAA AAATGTATCC ATGTGTGTCT GGAAGCATA TACACTTGTA 1020
CATTTTAA TACTCTGATT CTGTAACATT TCTGAGTTT GTTTGTGTT ACAGNAAAAA 1080
AAAAAAAGT GATAAAGCAA TCAGAAGACC AAGAGGTTA CTATGATGC TTAGGGTCGT 1140
50 CTGACCTTGG CTGGCCAATA GACCTACAG GCCAAATTAA TTTACGAGAG TAATAATTTT 1200
TCAAAAGCCA ATTTTCTTC TGTATTTCT GTATGAACT GCCAATATCA TGAATAGAAA 1260
55 GGGAGAACCA TAAAGGAGAA AGAACGTGAT GTTCTGTAT GTTCATGTAA ACCTAAAGAA 1320
ACAGTGTGGA GGCAGGCGC ATCAGCCGAA CTCTAGGAC TTGGTGTGTC TTGGAAGGCA 1380
TCCATACCTG CATTTTGCAT TCTTCGTATG TAATCATATT GCCAAAGACA AACTATTTCA 1440
60

TCATTTATTG TAAATAACAC TTTTCCCCAG ACCTACCATA AAGTTTCTGT GATGTATTGT 1500
CTTCCAGTTG CAATAAAAAAT TACTGAGTTG CATCAATTGA AGAAAAA AAAA 1560
5 CTCGA 1565

10 (2) INFORMATION FOR SEQ ID NO: 135:

(i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 2007 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 135:

20 TCTAAAGCC CCCTTATACC CCACTTTGTG CAGCAAAGAT CCCCGTGCAG GTCACAGCCT 60
GATTTGTGGC CAGGCTGGAC AAATTCCTGA GGCACAAC TT GGCTTCAGTT CAGATTTCAA 120
GCTGTGTGG TGTGGGACC AGCAGAAGGC AAACGTCCAG CCAACACACA GGACTGTAAG 180
25 AGGACTCTGA GCTACGTGCC CTGTGAAGAC CCCCAGGCTT TGTATAGGA GGTCTTCAG 240
CTTCCCCAAA GTCAGAGGTG ATTTGATTTG GGAAGACTG AATATTCACA CCTAAGTCGT 300
30 GAGCATATCC TGAGTTTAC TTCCTTATGG CTGCCCCTCC AAGTTCTCTC TCTCATACAC 360
ACACACACCC TTGCTCCAGA ATCACCAGAC ACCTCCATGG CTCCAGCTAT GGAACAGCT 420
GCATTGGGGC TGCCTTTCTG TTTGGCTTAG GAACTTCTGT GCTTCTTGTG GCTCCACTCG 480
35 CGAGGCAGCT CGGAGGTGTG GACTCCGATT GGGCTGCAGG CAGCTCTGGG ACGGCACAGG 540
GCGGGCGCTC TGATCAGCTC GTGTAAACA CACGCTCTTC TTGGCCTCCT GGCAGTTCTT 600
40 TCTGCGAATA GTCCTCTCCC TGGCCAGTTG AATGGGGGAA GCTGCTGGCA CAGGAAGGAG 660
AGGCGATCCC GGCTGAGGCT TAGGAAATTG CTGGAGCCGG CTCCAAGCAG ATAATTCCT 720
GGGAGGTTT TCAGAGTCAA ACATCATCTC GCCTGTCTTG GGGCCAGGT GTGTCACACA 780
45 AGCATCTCAA AGTCAAAAGC CATCTGGGGC TGCTGCTTCT CTCTCTCAGG CTCTGGGGAA 840
AGGAATCTCC CTCTCTCTC ACTTGATTC AAGTGTGGTT GAATGTCTG GAGCACTGGG 900
50 ACTTTTTCCT TCTTTTCCTT GATGGACCAA CAGTGCAAAT GCAATCTCGC CATTTAACTT 960
TCAGGTCGAT TTCCTTTCCT GATCAGACAT CTTTGTGCCC CCTTTAGGAA GGAAAAGAAT 1020
ACACCTACGA TGTGCCAGGC ACTGTGTTAG GCGCTTTTAT ATAGATCCTC GTTAGGATGA 1080
55 GACTAAGGGA TGAGGACATC TCTTTATAAA AGGCCCTTAA GTAATGGATA AACAGAAACA 1140
CTTAGAGGTG AGAAGGTCTG TCTCAAGAT CCAAGGTAAG ATTGCCTTCA GTCTGATGTT 1200
60 TGTCTCAAG GACTTATCCC CTACAATATT CTCCCACTCC ATACTTCTCC TTCTACCCCA 1260

CCATGTGCTC CCGTGCACTC CTCAGATGGT CAGAGGGGTA ACCCAAGTCC TTAGAGAATT 1320
TGGGGACCAA TAGAATATGT GATGTGTGAA TTTCTTTTAA AAAACTTAAG GAGTCTTTGC 1380
5 TACCTTCTGC TTGTTGAGTT GTTTTGCCAT TCATATTAAA AGCCAGCATC TCACTATTTA 1440
TTGACAGGTT GGGCTGTGTG TGTGCGCATG TGTGTATACA TTTCCAGGCG TGCCTGTGTC 1500
10 CTGTAGCTTT TTAAGAGGAA ACCCAGTCAT CCCACTATGA ATCTGGCATC TTCTTATGCT 1560
TCTAGTGTTC TGGCCATACA TCAACCAAGG GGTTTAATTT ATCCAATGCT TGACGACATG 1620
TTCAGGAGGG GCTGGATCAA ATTTTGAGAG GGTATGGA AAGGGAGGGG GAGAAGAAAT 1680
15 TGACATTTAT TTTATTATTT ATTTTAAATG TTTACATCTT CTTTATGTTG TATCAAGCCT 1740
GAATAGAAAC TGATAGCATT AAAATACTCC GTTCTCTCT CTCTTCTCGC TTCCTTTTTT 1800
20 TTTTTTTTTA AATTAGGAT AACACATTTT TGTTCCTAAA GTGATTGTG ATTTGTGCTG 1860
TATAAACTGT ATAAAAGGTT CTGTTTTTAA AGGTGGATT TCATTCCTCT GGGACAGTG 1920
GTCGCCAAGA CATCTACATT GTAAGAGAAC ACAGTGAAG ATCCTGTCCT GATTCTCAAA 1980
25 AATTATTTTC TCTGTATGAT TAAAAGT 2007

30

(2) INFORMATION FOR SEQ ID NO: 136:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1291 base pairs
35 (B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 136:

CTTTTAACCC TCCCCCTCA CACACATACA TATCAGGTG TTTTCTAGT AAAAACCCAA 60
GTAGCTCAGA TTCTACTTTA ATGTCAGTGC AGATTGTCAT TGAATCATGC CATTATGTTT 120
45 TTTCTCATTT TTATGCTGTT GGTCTTAGT TTTTAAATG ATATAAGAA CTCAGCAATG 180
GTTTATTTTT CTAATCATAC TTAGGGTTTA GGAAACACTA CCACTAGTTA TCATTTAATC 240
AATCTCAATG GTCTACTGAA ACAAAAATGG TAACTTTTCA TTAGTGGATT ATTTAGAGTT 300
50 ATAGTAGTTG TTTCCAGAAA AACTTCCTC ACAATTGTAC TTCCCAATCA AATCATGTGA 360
TCATACAGTT ATTCCCATGA AAGGCAGAAT GTTTGTTTCA AAATTAATCT AGTTTTCTGT 420
55 ACATTTAAAT TTGAGAAGGT GACAACTGGC TCTTTTCCAG TCTTCCTTCA TGTAGTTTTT 480
CTGATAGACC ACTATTGGCA AACAGTATCT GTCAACTACC AAATGTGTAA AATTTTCTGT 540
ATTCACATTT GTCTTATTG TAAATAGTGA ACTAAAATTT TTGGCAGATC AGCAACATTT 600
60

GCTGAGCCTG TTTTTTAAGC TAATGTGTAT TCTTACTAAT GTTCCTATCA AGAATGGATT 660
TGTAATATAT GCTGTCTATT TCTAATGTTC ACATTCATAT TTTGAGGTTC TATCTTATTT 720
5 TAATAGAGAA CAGACTTCTC AAAAAATCTT CAGAAGCAGC TTATTATTGA AATATCGAAA 780
TATTGAAATA AACCCGGTGG GTTAGATTAC TCATCTGTCC ACCAAGTGGG ACATTTCAT 840
GGACTGGGGG CTTAAAGGAC TTAGAAGAGA CCTGTAAGTA AATCCTGAAA ATGAGCCAAT 900
10 CCCCACTTGA ATGGTTACTG GAGTAAACCC ACCTTTACCA CCCCAATTAC AGCACCCGAG 960
GCCGATAAAC CAACTTGGCT CTGGTTCATT TTTCTTTTCT TCATTGTGA TGCTCAGATT 1020
15 CAAAAATGTGT GTTCTACACT GTTACAGGCT TCTCTTTTGT TTGATTAAAG ATTTTAGTCC 1080
TACTTTTGTA TGGACACATT AGAATATTCA GAGACCAAAA TAGAAGAATT TGCTGTTAGA 1140
TATTTTTCAG AAGTCAGCAG ATTGTGGCA AATCATTTAT TTGCCTTTT AAAAATTCAT 1200
20 TTAAGCAGTT CAGAGAGTAG ACTACTCAGA AAATTATTTT ACGTAATTGT CTAAGAGGTC 1260
AATATTTTTT AATGCATATT GAATCAAATA A 1291
25

(2) INFORMATION FOR SEQ ID NO: 137:

30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1906 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 137:

GGCACGAGGA CCTACTTTTG TAACAGACCA TGGTTGTGTC CAAGGTAAAA CCACAGTGAT 60
40 ATTTTGGAT GCTTTGTCTG CAATCTTGAC TTGTTTTCG AGTATCATTA TTCAGACTTC 120
AAATTGTGAA TCTTTTAAAC ATCTTGATAA TTTGTTGTG AGAGCTGTTC ATTCTAAAAT 180
GTAATGAAAT TCAGTCTAGT TCTGCTGATA AAGATCATCA GTTTTGAAAG GTTACTGATT 240
45 TTCTCTTCC CTCTTAGTTT TTTACCCAAT ATATGGAGAA GAGTAATGGT CAATCTTAAC 300
ATTTGTTTT AATGTTTAA TAAAGCTGCT GGGCAGTGGT GCAGCATTC TACCTAGTGT 360
50 CATAAAAGCA AAATACTTAC ATAGCTTTCT TAAATATAG GAATGACATT ACATTTTTCAG 420
GAGAAAGTAA GTTGCTTTGC ACCGCCTACT TAATTCCTTT CCATATATTG TGATACAAAC 480
TTTTGAATAT GGAATCTTAC TATTTGAATA GAAATGTGTA TGTATAATAT ACATACATAC 540
55 ATAAGCATAT ATGTGTGTGT GTGTGTGTAT ATATATATAT ATGCATGCTG TGAAACTTGA 600
CTACACAACA TAAATCACTT TTTAAATTC AGGAACGGGT AGTCTGACAC GGTGATTATC 660
60 CTTTGTAGGC TGAATCCGTT ATTAACCTGT TATTTAGGTT TTAATCCAG TAGCAAGGGA 720

TTCTAAGTTA GTTGCACTTA CATGATTATT GTGATTTAAA ACTAAGAATA AAGGCTGCAT 780
TTTCAAAGAT AAATGGAAT TGCTGTTGGT GAAATAACAA CCAAATACT GAATCTGATG 840
5 TACATACAGG TTTCTACAGG AAGAGATGGT ATAATTTACA ATTTGGAGAT TTAATAACCA 900
GGGCTACCCA GAAAAAGTGA CTTGATAACA TGGTACCAAT AAGTAAGGGA TGCTCTCTCG 960
10 GTTGTCTTTT GCCACTTTCA AGATTTTAAC TTCTCAGGTT ATTAATCAAA ATTATTGTAT 1020
AAGTTAGCCA ATAGAATTTT TAGGTTAAAA CAACAGATGG GGGGTTTGTG GAGTGTTTAA 1080
TGTCATGGGC ATTTTITAGTA GCATAGACCC TTGTCTCTGC ATTTGAATGT TTCGTATATT 1140
15 TTTGTTTCAC AGTTAATCTT CCCTCCCCAA GTTTGCTATT CAAATCAACT GCCTGAATGA 1200
CATTTCTAGT AGTCTGATGT ATTTTCTGA GGAATAGTTT GTGATTCCAA TGCAGGTGTC 1260
20 TTCATTACCA TTACCTCTAC ACTGCAGAAG AAGCAAACT CCTTTATTAG AATTACTGCA 1320
CATGTGTATG GGGAAAATAG TTCTGAAAGG CTAGAATGAT ACAAGTGAGC AAAAGTTGGT 1380
CAGCTTGGCT ATGGAGTGGT GGCAATAATC TCTAAACATT CCAAAGACC ATGAGCTGAA 1440
25 CCTAAACTCC CTTGGGAATC TGAACAAAG GAATATGAAA ATTGCCATTT GAAACTGAC 1500
CAGCTAATCT GGACCTCAGA GATAGATCAG CCAGTGGCCC AAAGCCATTT CAAGTACAGA 1560
30 AATTATAGAG ACTACAGCTA AATAAATTG AACATTAAAT ATAATTTTAC CACTTTTGT 1620
CTTTATAAGC ATATTGTAA ACTCAGAACT GAGCAGAACT GACTTTACTT TCTCAAGTTT 1680
GATACTGAGT TGACTGTTC CTTATCCCTC ACCCTTCCCC TTCCCTTTCC TAAGGCAATA 1740
35 GTGCACAACT TAGGTTATTT TTGCTTCCGA ATTTGAATGA AAACTTAAT GCCATGGATT 1800
TTTTTCTTTT GCAAGACACC TGTTTATCAT CTTGTTTAAA TGTAATGTC CCCTTATGCT 1860
40 TTTGAAATAA ATTTCTTTT GTAAAAAAA AAAAAAAA AAAAAA 1906

45 (2) INFORMATION FOR SEQ ID NO: 138:

(i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 1935 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 138:

55 TCTGAATAA TGCTAACAGA TCCCCCTGAG GGATTCTTGA TGGGCTGAGC AGCTGGCTGG 60
AGCTAGTACT GACTGACATT CATTGTGATG AGGGCAGCTT TCTGGTACAG GATTCTAAGC 120
TCTATGTTTT ATATACATTT TCATCTGTAC TTGCACCTCA CTTTACACAA GAGGAACTA 180
60

	TGCAAAGTTA GCTGGATCGC TCAAGGTCAC TTAGGTAAGT TGGCAAGTCC ATGCTTCCCA	240
	CTCAGCTCCT CAGGTCAGCA AGTCTACTTC TCTGCCTATT TTGTATACTC TCTTTAATAT	300
5	GTGCCCTAGCT TTGGAAAGTC TAGAATGGGT CCCTGGTGCY TTTTTFACITTT GAAGAAATCA	360
	GTTTCTGCCT CTMTTGGAA AAGAAAACAA AGTGCAATTG TTTTTFACITG GAAAGTTACC	420
10	CAATAGCATG AGGTGAACAG GACGTAGTTN AGGCCTTCCT GTAAACAGAA AATCATATCA	480
	AAACACTATC TTCCCATCTG TTTCTCAATG CCTGCTACTT CTGTAGATA TTTCAITTTCA	540
	GGAGAGCAGC AGTTAAACCC GTGGATTTTG TAGITAGGAA CCTGGGKTCA AACCTCTTC	600
15	CACTAATTGG CTATGTCTCT GGACAAGTTT TTTTTFITTTT TTTTTFITTA ACCCTTCTG	660
	AACITTCACIT TTTATGTCT ACCTCAAAGA ATTGTGTGA GGCTTGAGAT AATGCATTTG	720
20	TAAAGGGTCT GCCAGATAGG AAGATGCTAG TTATGGATTT ACAAGGTGT TAAGGCTGTA	780
	AGAGITCAA ACCTACAGTG AATCACAATG CATTTACCCC CACTGACTTG GACATAAGTG	840
	AAACTAGCC AGAAGTCTCT TTTTCAAAT ACTTACAGGT TATTCATAT AAAATTTTIG	900
25	TAATGGATAA TCTTATTTAT CTAAACTAAA GCTTCCTGTT TATACACACT CCTGTTATTC	960
	TGGGATAAGA TAAATGACCA CAGTACCTTA ATTTCTAGGT GGTGCTCTGT GATGGTTCAT	1020
30	TGTAGGTAAG GACATTTTCT YTTTTTCAGC AGCTGTGTAG GTCCAGAGCC TCTGGGAGAG	1080
	GAGGGGGGTA GCATGCACCC AGCAGGGGAC TGAAGTGGGA AACTCAAGGT TCTTTTACT	1140
	GTGGGGTAGT GAGCTGCCTT TCTGTGATCG GTTTCCTAG GGATGTTGCT GTTCCCCTCC	1200
35	TTGCTATTCG CAGCTACATA CAACGTGGCC AACCCAGTA GGCTGATCCT ATATATGATC	1260
	AGTGCTGGTG CTGACTCTCA ATAGCCCCAC CCAAGCTGGC TATAGGTTTA CAGATACATT	1320
40	AATTAGGCAA CCTAAAATAT TGATGCTGGT GTTGGTGTGA CATAATGCTA TGGCCAGAAC	1380
	TGAAACTTAG AGTTATAATT CATGTATTAG GGTTCCTCCAG AGGGACAGAA TTAGTAGGAT	1440
	ATATGTATAT ATGAAAGGA GGTATTAGG GAGAACTGGC TCCCACAGTT AGAAGGCGAA	1500
45	GTGCGACAAT AGGCCGTCTG CAAGCTGGGT TAGAGAGAAG CCAGTAGTGG CTCAGCCTGA	1560
	GTTCAAAAAC CTCAAACTG GGAAGCTGA CAGTGCAGCC AGCCTTCAGT CTGTGGCCAA	1620
50	AGGCCAAGAG CCCCTGGCAA CCAACCCACT GTGCAAGTC CTAGATTCCA AAGGCTGAAG	1680
	AACCTGGAGT CTGATGTCCA AGAGCAGGAA GAGTGAAGA AAGCCAGAAG ACTCAGCAA	1740
	CAAGGTAGAC AGTGTCTACC ACCAYAGTGG CCATACCAA GAGGCTACCG ATTCTTCCT	1800
55	GCTACCTGGA TCCCTGAAGT TGCCCTGGTC TCTGCACCTT CTAAACCTAG TTCTTAAGAG	1860
	CTTTCCATTA CATGAGCTGT CTCAAAGCCC TCCAATWAAT TCTCAGTGA AGYTTCAAAA	1920
60	AAAAAAAAA AAAAA	1935

(2) INFORMATION FOR SEQ ID NO: 139:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1446 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 139:

5	NGCCCCCTTG GCACAAGTCA GATGAAGCAC GTTCTGCCGG GGAGGCCCTC AMCTTCCAGA	60
15	GAGGACAGAC ACAGATTTCG TGCTGGGGGA GGGAGGAGTC CACGCATCCT GATGCTGCCT	120
	GGAAGCTTAT TTTCCCGTGG CCAGGATGCA TTTCTCTGAG TGGAAACAGG TTCTTGCA TG	180
20	TGGATGTGTG TTTCCCCAGG CAGACGGCCC CTCTYTTCCC AGCACTTCCC TGCCTCCCCC	240
	AGGCCTCAGG CCAGCACCCA GTTCCTCCTC ACATGGCAGG TGAGCACAGA CTTCTAGTTG	300
25	GCAGGAGCTG AGGAGGGTGA ACAAAACCCG AGGGAGGCCG GGCCTTGCT CCCGAGTTGG	360
	GGGGAGGGGG TGTGGCAACG TGCCCCCGC AGAGGCCACG CATGTTTGAC CAAAGCCCTC	420
	ATTGTGGTCC GAGGACAGCC TTTTCCCCAG GCCTCARAGC ATTGCTCATC CGTGCCAAAC	480
30	TGGGTAGGTG GATTTGAGCG GAAAGACTCC CAAATGTGC CAAGAATTTC CCRGTCCAG	540
	GCAGGGCAGG GGAAACTAAG GGCAAGCAGG ATACAGGGCG AGGGATGTGG CAGGTGAGGG	600
	GGCTCCCGCC TGTGCCCTT CTCTCACCA TGTCTCCCC ACCCTGCCTC AGTTCTCCGT	660
35	TCCCCCTCAT CTCGTCCCC CTCTTTGAAG CTGTCCCAT CTCAGTGTC GACCAGCCTT	720
	CTCTCAKCT GACCACCCTC CTCTGACCSA CGCCCCCTCC TTGTCTGAAA AAAGGAGCCT	780
40	TGAATGGTGG AGGGAGGCAG TGGGGAGAAA GGTCTCACCG GACAGGTGG GAGAATGAGG	840
	TCAGCGGTGC TGGGGAACAG ATGGAGGGG CAGTGGGAC AGGGCTTGGG CAGACACCAG	900
	CAGGAATAAT TTGAAATGTG TGAGGTGACT CCCCGGAGGC CTTGGGCTTG GGCATTGCGG	960
45	AAAAGAATGA TGTCTGGAAG GGCTTAAGG ACACAGTGA CGAGGGGAGA GTCCTCATCT	1020
	GCTGGCATTT TGTGGGTGT TAGTGCCAAA CTTGAATAGG GGCTGGGTG CTGTCTTCCA	1080
50	CTGACACCCA AATCCAGAAT CCCTGGTCTT GAGTCCCAG AACTTTGCCT CTTGACTGTC	1140
	CCTTCTCTC CTACCTCCAT CCATGGAAAA TTAGTTATIT TCTGATCCTT TCCCCTGCCT	1200
	GGTCTAGCTC CTCTCCAAAC AGCCATGCCC TCCAAATGCT AGAGACCTGG GCCCTGAACC	1260
55	CTGTAGACAG ATGCCCTCAG AATTGGGGCA TGGGAGGGG GSTGGGGGAC CCCATGATTC	1320
	AGCCACGAC TCCAATGCCC AGCTCCTCTC CCCAAAACAA TCCCACAAAT CCCTTATCCC	1380
60	TACCCCAACC CTTTGGCGCT CTGTACACAT TTTTAAACCT GGCAAAAGAT GAAGAGAATA	1440

TTGTAA

1446

5

(2) INFORMATION FOR SEQ ID NO: 140:

- 10 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1109 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 140:

TTTTTTTTTT TTTGATATGA AATTGCTTTT CTCCATTGCA GAAATAAGCT AGGGAAACAC 60
TAACCCAAAA ACTTTCTGTA GAGCTGTTCC TTTGGAGGCA GCATCACTTA TTGGCAGTAA 120
20 AGACTCAGTA TAAAAGCACC AGCATCCCTA CTGGGTGAT GGGGATTAAT TTTATAGCAT 180
TCCATTTTCC TAGTGCCACA TGTGAAATG GATTTTGATG ATCTTAATCT ATATTCTACC 240
25 CTTATAATAA AAGATCAAAA GATATATCTC CTATGAACAG ATGGAGATA GGAGATGAAA 300
AGTTGGGAGG ATGCTTTTAT TCTAATGTGA GGTAGGGAA AATGTGGATA ACATTACTGG 360
GGTGARGGAG GCATGTGTTCT TTAGTTGGAG TTCTCATTTT TATTCTCCAG TACTGACTTG 420
30 TGGGGAAAGC ATACTTTTTC ACTGCCAGGT ACTGAATGCA GAGGCTCAGT GAAGTATATA 480
TGTGGGAAGT GCATGCATTT CGTTTATTAG CAAACATAGC TGGATTAAGA CAAAGTTGTT 540
35 GGTTTGGAAA GGGGTAAAG CCTTAAGTGA ACAAATCTAG CTAACAGTGA ATGAACTAGG 600
TAATATAACT TGCATATTTT TAATTTCTTT TGGTTAAAG TCCCCATAC TTCTCTGTTT 660
GGAGACATGA GAAGTATGAT TACTTCAGTG TTAGTTTCTT TAATTTTTTT TTTCCCCTAT 720
40 TTGTCCCTTG TCACTTTGTT GCAAGCTAGA AATCTGTGGG TTATACATAG GGCAGCTCTT 780
TGTGAAAGTG GTTTATTCCA CTGGAGAAAG GGGATTGAAA ATCAGTTAGA ACCAATGTAT 840
45 TTCTTGCCCC ACGGAACACT ATTCTATATA GATAGCTGAA AGAAGCTGCT GTGAGGAGCT 900
CAGCTCCAAA CACAGGATCA GCACCTTGTA TAGGAATTCC CATGAATTAT GACTTCTCAT 960
TCTGTTTTAT CAGAGTGCAT ATATGTCCTA CTTCAGGAAA AGTAAACAG TCATTTACGA 1020
50 AAGAAAGTCA ATCTGTATCC TAAGCATTTT AATAAAAAGT TAAAACAAAA AATTAAAGG 1080
GACACTCGAG GGGGGGCCCG AAACCCAAT 1109

55

(2) INFORMATION FOR SEQ ID NO: 141:

60 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 497 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 141:

TAGGACTAAC TTAAATTCTT TTATTCATCT TTTATTTATT AAAAAATTTT ATTTCTTTGA 60

10 ATTTTCCTGT AATTTCCTTA RGCTCTCTA TAAATGTTA TATTCATGTG AACCATACCT 120

CATTATCCTT AACATTTACT CTCAAAAAGC TTTTATTTT TATTTTTTTG AAGGTAGTTT 180

15 TTCTGTGTGT ACTCTGTAAC ATGATTTTGC TTTCAAATCA TTGTGTGTC CCCATACAAA 240

ATGCCTTTTA TTTTGTAGGA TCGTGGACTT TTTAGTATGG CATGAGTGTG CTAAAAGCCA 300

GATATCTTTC CACATTCACT GGTGGCTTTG ACACCTAGTT TTTAATCTCC CATCCTTACT 360

20 TTAAACCCCTG ACAGTGCAGT CCTCAGTCAG GGCCAGGACC GGGCTGAGGC CCTTGTGGA 420

GATGCTGCAC CACCAGCAGA AGGCTGAGAC CTGGTTACCT GTACCTGTTC ACTTGTAATA 480

25 AAAAGAATTA TCTAAAA 497

30

(2) INFORMATION FOR SEQ ID NO: 142:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 269 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 142:

ATGAGGCAGA GGCAAGCTGC CTGCCAAGCC CCTCCCTCAA GGAATGGCCT TGCCAGGAA 60

40 TGCCACCAC ACATACCCTC TTCTTTTTTT CTAGTCAAAC TCTGTATTAT TCCTGGCTT 120

GCCTCCCTCC TTCTCTCCC TCTCAACCTT TTAATTCTGG TTTCTATTTC ATGGGATTTG 180

45 GGGTTGAAGT TAAACTTACA ACAGTGCCGC CAACACCAAG TCTTGAGGA AAAAAATACA 240

AAGAAATTTA AAAAAAAAAA AAAAAAAAAA 269

50

(2) INFORMATION FOR SEQ ID NO: 143:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1269 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

55

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 143:

TTGATTGACT ATGGTCTCTC CGGCTACCAG GAAGAGTCTG CCGAAGTGAA GGCCATGGAC 60
TTCATCACCT CCACAGCCAT CCTGCCCTG CTGTTCGGCT GCCTGGGCGT CTTCCGGCCTC 120
5 TTCCGGCTGC TGCAGTGGGT GCGCGGAAG GCCTACCTGC GGAATGCTGT GGTGGTGATC 180
ACAGGCGCCA CCTCAGGGCT GGGCAAAGAA TGTGCAAAAG TCTTCTATGC TCGGGTGCT 240
10 AAACTGGTGC TCTGTGCCG GAATGGTGG GCCCTAGAAG AGCTCATCAG AGAACTCACC 300
GCTTCTCATG CCACCAAGGT GCAGACACAC AAGCCTTACT TGGTGACCTT CGACCTCACA 360
GACTCTGGGG CCATAGTTGC AGCAGCAGCT GAGATCCTGC AGTGCTTTGG CTATGTCGAC 420
15 ATACTTGTCA ACAATGCTGG GATCAGCTAC CGTGGTACCA TCATGGACAC CACAGTGGAT 480
GTGGACAAGA GGTTCATGGA GACAACTAC TTTGGCCCAG TTGCTCTAAC GAAAGCACTC 540
20 CTGCCCCTCA TGATCAAGAG GAGGCAAGGC CACATTGTCG CCATCAGCAG CATCCAGGGC 600
AAGATGAGCA TTCCTTTTCG ATCAGCATAT GCAGCCTCCA AGCACGCAAC CCAGGCTTTC 660
TTTGACTGTC TCGTGCCGA GATGGAACAG TATGAAATTG AGGTGACCGT CATCAGCCCC 720
25 GGCTACATCC ACACCAACCT CTCTGTAAAT GCCATCACCG CGGATGGATC TAGGTATGGA 780
GTTATGGACA CCACCACAGC CCAGGGCCGA AGCCCTGTGG AGGTGGCCCA GGATGTTCTT 840
30 GCTGCTGTGG GGAAGAAGAA GAAAGATGTG ATCCTGGCTG ACTTACTGCC TTCCTTGGCT 900
GTTTATCTTC GAACTCTGGC TCCTGGGCTC TTCTTCAGCC TCATGCCTCC AGGGCCAGAA 960
AAGAGCGGAA ATCCAAGAAC TCCTAGTACT CTGACCAGCC AGGGCCAGGG CAGAGAAGCA 1020
35 GCACTCTTAG GCTTGCTTAC TCTACAAGG ACAGTTGCAT TTGTTGAGAC TTTAATGGAG 1080
ATTTGTCTCA CAAGTGGGAA AGACTGAAGA AACACATCTC GTGCAGATCT GCTGGCAGAG 1140
40 GACAATCAAA AACGACAACA AGCTTCTTCC CAGGGTGAGG GGAAACACTT AAGGAATAAA 1200
TATGGAGCTG GGGTTTAAAC CTAAAACTA GAAATAAACA TCTCAAACAG TAAAAAATAA 1260
AAAAAAAC 1269
45

(2) INFORMATION FOR SEQ ID NO: 144:

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1944 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

55

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 144:

AAAAGGCAAA CTATAGGATA ACACAGAGCC CTTTGTGAAA ATAAATTGGC ATTGGAGTGT 60
60

	TTTACCCCTCT AGCTGTTTTA CTTAGAATGT AACATATGCT GCCTACCCAC CTCAAAATGT	120
	CTGTACTGCA AGAGGGCCCT GGGCCTCTGC TTTCCATATT CACGTTTGGC CAGAGTTGTA	180
5	GTCCCAAAGA AGAGCATGGG TGGCAGATGG TAGGGAATTG AACTGGCCTG TGCAATGGGC	240
	ATGGAGCACA AGGGGTCACA GCATGCCTCC TGCCCTACCG TGGCAGTACG GAGACAGTCC	300
	AGAACATGGT CTTCTTGCCA CGGGGTGTG TGTCTCTCG TGGTCTGCA TGTCTGTGGC	360
10	TCACCTTTAT TCTTGAACT GAGGTTTACC TGGATCTGGC TACTGAGGCT AGAGCCCACA	420
	GCAGAATGGG GTTGGCCTG TGGCCCCAA ACTAGGGGGT GTGGGTCAT CACAGTGTG	480
15	CCTTTTGTCT CCTAAAGATA GGGATCTACT TTTGAAGGA ATTGTTCCCT CCAAATAAAT	540
	TGTCTTACC TTGGTCCCTT CTTTGTGCC AGTATTCAAG TGGTATAGCT CTGAGCAGGG	600
	TCACATTTGG CCAAACCTGA CACTGTCTTG CTGCATCTC CTTTGGCAA CATCAGGGTC	660
20	AGAAITCAGG ATAGCCCTTC CTAGGGCACT GGACTTTCTG GCATGGGGG TGTGTTTGCA	720
	CAAGTTATTT TCATGTTACC TGGAGAGTGT CCAGAGGCTG CTCTGAGGCT GAGGTGTGTT	780
25	CCCCCTTGCC TGGTTCAGC TGTGAGGG ATACCATCCT AGGGTCTGGG AATCCAAGGC	840
	CACGAGACTC CTTGGTTTGT GGTCCGAGAT CCTGTACTAA GGAGGGTCTG GCCAGAGGAA	900
	CAGACCAGCT TTTGCACAAT GAAGCGCAAG GGAACAAGTG GTTTGCCTGG TGTCTTACCT	960
30	GTCCTGAACC TGGTCTGTG GGCCATTGAA AAGTTAGATC TGTGATCTCT GGGGTTTTTG	1020
	TGGCTTTGTT CAATGCTTCC ACTCTAGGGC AGGCAGAGCA GTCTATACTC TCCCAAGCCT	1080
35	GCTTGACCTC CAAGTAGAGC TGATACAGAG ATCTGTGAAT ATTGTGATAG AAATTCTTTG	1140
	GTATTCATAC ATTTTCACTG CAAGTCAGCA ATTTCCAGG TACCATGTAA GCTATAAAAC	1200
	AGTCATTTCT AAAGACAGAG GATAGCTGTG ACTCATGGGA TCATGAGGTC CATGGCTGGT	1260
40	TGCAGGTGCC CTTTTCCTT CTTGAGGTTT TGTCTCTTCC TGTGTTGTCC CCAGCAAGGG	1320
	AGAGACTGTG GGGTGGATTG GGAGAACAGA TTAGGAGTAT AGCAAATGAA CCCAGAATGG	1380
45	AACAGTGGGG AGCTAACTGT GAATGAGGAG AGTACCTGCT GCAGGACCTG GAGGTCAGGT	1440
	GTGAATGCTG TATTGGCACA GGAATAAAT ATCCTGGCGT CTGGAGCCTT CACCTCTCCG	1500
	TCAAGTCCCT CCTGTGATAC TGCCATGGCA CAGGATCTGA GTTGCAGCTC TGCACCCTAA	1560
50	ATCACACCTT GGCATGTGTC TGGGCTGCAG GGCTGCCAGG TTCTGTACTT GTGTCCAGCT	1620
	GTGGCCCTGG ATGCTGGAGC TGGAGGGTTT TCTGTGCTCA GACTGTAGCC TGTAGCTCTT	1680
55	GGCCTGTGTA GAGCCCCCTC CTGTGCCCTC AGTGGCTGTC GTTGTGTAAC ATCATCAGGA	1740
	AGATGGGAAA GGTGAGCAG AATTTTCTG CCCTACAAAG GGTGGAAGAG AAAGGACACA	1800
60	GTATTTTCAT GAATTACCA TATATCTTTG TTTTCTTCA ACGAAAAAGT TAATTGAGGC	1860

AATGTCATCT GCTCAAAGTT GAGTGGTTTA TTCACAATAA ACTGTAAGTT TCTGATTATA 1920
AAAAAAAAAA AAAAAAAAAA AAAG 1944

5

(2) INFORMATION FOR SEQ ID NO: 145:

10

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1021 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 145:

TCGACCCACG CGTCCGGGGT GCGCAACGGG GAGTTCCGGC TGGAGACCCG TGCTCTGGGC 60
20 CGGCGCCTTC ACCATGGCCT CGGCAGAGCT GGAATACACC ATCGAGATCC CGGATCAGCC 120
CTGCTGGAGC CAGAAGAACA GCGCCAGCCC AGGTGGGAAG GAGGCAGAAA CTCGGCAGCC 180
TGTGGTGATT CTYTTGGGCT GGGGTGGCTG CAAGGACAAG AACCTTGCCA AGTACAGTGC 240
25 CATCTACCAC AAAAGGGGCT GCATCGTAAT CCGATACACA GCGCCGTGGC ACATGGTCTT 300
CTTCTCCGAG TCACTGGGTA TCCCTTCACT TCGTGTTTTG GCGCAGAAGC TGCTCGAGCT 360
30 GCTCTTTGAT TATGAGATTG AGAAGGAGCC CCTGCTCTTC CATGCTTCA GCAACGGTGG 420
CGTCATGCTG TACCGCTACG TGCTGGAGCT CCTGCAGACC CGTCGCTTCT GCGGCCTGCG 480
TGTGGTGGGC ACCATCTTTG ACAGCGCTCC TGGTGACAGC AACCTGGTAG GGGCTCTGCG 540
35 GCGCCTGGCA GCCATCCTGG AGCGCCGGGC CGCCATGCTG CGCCTGTTGC TGCTGGTGGC 600
CTTTGCCCTG GTGGTCGTCC TGTTCACGT CCTGCTTGCT CCCATCACAG CCNTCTTCCA 660
40 CACCCACTTC TATGACAGGC TACAGGACGC GGGCTCTGCG TGGCCGAGC TCTACCTCTA 720
CTCGAGGGCT GACGAAGTAG TCCTGGCCAG AGACATAGAA CGCATGGTGG AGGCACGCCT 780
GGCAGCCCGG GTCCTGGCGC GTTCTGTGGA TTTCTGTGTA TCTGCACACG TCAGCCACCT 840
45 CCGTGACTAC CCTACTTACT ACACAAGCCT CTGTGTCGAC TTCATGCGCA ACTGCGTCCG 900
CTGCTGAGGC CATGCTCCA TCTCACCTCT GCTCCAGAAA TAAATGCCTG ACACCTCCCC 960
50 ACAAAAAAAAA AAAAAAAAAA ACTCGAGGGG GGGCCCGGTA CCCAATTCGC CCTATAAAGG 1020
T 1021

55

(2) INFORMATION FOR SEQ ID NO: 146:

60

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1285 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 146:

GGCAGAGGA GGGCCACGGC AGCCATCGCG CTTTGCAGTT CGGTCTCCTG GTGTACGGCC 60
AACGCCAAGT AGGGGATTGC GTTCCCTCCA GTCGCAGACC CTATCAGATT TGGATATGTC 120
10 CTTTCATATTT GATTGGATT TACAGTGGTT CAGCAGTGTG CTACAGTTTT TAGGATTATA 180
TAAGAAACT GGTAACTGG TATTTCTTGG ATTGGATAAT GCAGGAAAA CAACATTGCT 240
15 ACACATGCTA AAAGATGACA GACTTGGACA ACATGTCCCA ACATTACATC CCAGTTCCGA 300
AGAACTGACC ATTGCTGGCA TGACGTTTAC AACTTTTGAT CTGGGTGGAC ATGTTCAAGC 360
TCGAAGAGTG TGGAAAACT ACCTTCCTGC TATCAATGGC ATTGTATTTT TGGTGGATTG 420
20 TGCAGACCAC GAAAGGCTGT TAGAGTCAAA AGAAGAACTT GATTACTAA TGACAGATGA 480
AACCATTGCT AATGTGCCTA TACTGATTCT TGGGAATAAG ATCGACAGAC CTGAAGCCAT 540
25 CAGTGAAGAG AGGTTGCGAG AGATGTTTGG TTTATATGGT CAGACAACAG GAAAGGGGAG 600
TATATCTCTG AAAGAACTGA ATGCCCGACC CTTAGAAGTT TTCATGTGTA GTGTGCTCAA 660
AAGACAAGGT TACGGAGAAG GCTTCCGCTG GATGGCACAG TACATTGATT AACACAACT 720
30 CACATGGTT CCAGGTCTCA ACGTTCAGGC TTAATCAGAG ATTGATTGC TCAACATGCA 780
TAACTTGAAT TCAATAGACT TTTGCTGGTT ATAAACAGA TGTTTTTTAG ATTATTAATA 840
35 TTAAATCAAC TTAATTGAA TGAGAATTGA AACTGATTC AAGTAAGTTT GAGTATCACA 900
ATGTTAGCTT TCTAATCCA TAAAGTACT TGGTTTTTAC AGTTTATAAT CTGACATCAC 960
CCCAGCGCCA TTTGTAAAGA GCAACTTTCC AGCAGTACAT TTGAAGCACT TTTTAACAAC 1020
40 ATGAACTAT AAACCATATT TAAAGCTCA TCATGTTAAA TTTTTATGT ACTTTTCTGG 1080
AACTAGTTTT TAAATTTAG ATTATATGTC CACCTATCKT AAGTGACAG TTAATAATTA 1140
45 GCTTATCAA TGATTGCATG ATGCCTTACA GTTTTCAATA ACTTTTTTTC TTATGCAAAC 1200
GTCATGCAAT AAAACAACT CTAATGTTG GCAAAAAAAA AAAAAAAAAA NTCGAGGGGG 1260
50 GGCCCGTACC CAATTCGCC TAAAG 1285

55 (2) INFORMATION FOR SEQ ID NO: 147:

55

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1386 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
60 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 147:

5 GGCACGAGGT GCGCAGGGG TCAGTGGTTC TCTCGGGTCT CGGGACAGGT GAGCACCCCTG 60
ATGAAGGCCA CGGTCCTGAT GCGGCACCTG GCGGGGTGCA GGAGATCGTG GCGGCCCTCC 120
GCAAGGGCGS CGGAGACCGG TTACAGGTGA TTTCTGATTT TRACATGACC TTGAGCAGGT 180
10 TTGCATATAA TGGAAAGCGA TGCCCTTCTT CTTACAATAT TCTGGATAAT AGCAAGATCA 240
TCAGTGAGGA GTGTCGAAA GAGCTCACAG CGCTCCTTCA CCACTATTAC CCAATTGAGA 300
TCGACCCACA CCGGACCGTC AAGGAGAAGC TACCTCATAT GGTGGAATGG TGGACCAAAG 360
15 CGCACAATCT CCTATGTCAG CAGAAGATTC AGAAGTTTCA GATAGCCCAG GTGGTTAGAG 420
AGTCCAATGC AATGCTCAGG GAGGGATATA AGACCTTCTT CAACACACTC TACCATAACA 480
20 ACATCCCCCT TTTCATCTTT TCTCGGGCA TTGGTGATAT CCTGGAAGAA ATTATCCGAC 540
AGATGAAAGT GTTCCACCCC AACATCCACA TCGTGTCTAA CTACATGGAT TTTAATGAAG 600
ATGGTTTTCT CCAGGGATTT AAGGGCCAGC TGATACACAC ATACAACAAG AACAGCTCTG 660
25 TGTTGTGAGAA CTSTGGTTAC TTCCAGCAAC TTGAGGGCAA AACCAATGTC ATCCTGCTGG 720
GAGACTCTAT CGGGACCTC ACCATGGCCG ATGGGGITCC TGGTGTGCAG AACATTCTCA 780
30 AAAITGGCTT CCTGAATGAC AAGGTGGAGG AGCGGCGGGA NCGCTACATG GACTCCTATG 840
ACATCGTGCT GGAGAAGGAC GAGACTCTGG ATGTGGTCAA CGGGCTACTG CAGCACATCC 900
TGTTGCCAGG GGTCCAGCTG GAGATGCAAG GCCCCGAAG GCGCAGGCTN CCAGNCCGCC 960
35 TGCAGGCCGT GGTGAGGAGG GCGCCTCCC CAGAGTCTGC TCCCCCGTGA ACACAGAGCA 1020
GANGCCAGG TGGCCAGCAG TGGCTGGGTC CTTCCGCGCC CCTCCGTCTT CCTTTCCTG 1080
40 AGCACCTTCA TCACCAGAGG CTTGAAGGAA CCCCAGCATG TGGCAGGCA CAGGCACTGT 1140
TCCTGGTGAA CCTTGGACCA CAGCATGTCA GTGCTCTAGG GATTGTCTAC TCCAGGGATT 1200
TTCTTCAAAA TTTTAAACA TGGGAAGTTC AAACAAATAT AATGTGTGAA ACAGATCAAA 1260
45 ATTTTAAAA TGAAAAAAA GCTGCTCTGA TTCAGGGGAT GTGGGTCGGG GTAGAACCTG 1320
GACCTCTTGG CCTGGGGCA CATGGGATGC TTCTAGGAAC ACAGTTTGAG AACCACCAA 1380
50 AAAAAA 1386

55 (2) INFORMATION FOR SEQ ID NO: 148:

(i) SEQUENCE CHARACTERISTICS:

- 60 (A) LENGTH: 2098 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 148:

5	AGCCCTTCTC CCCGCGCTTG GGA CTCTGAC ATCTTAAGGC TGCACGGTCG TGTCTTGTC	60
	TGGGTGAGGC CATGTCGTG ATCCAAGGTT CCTGGAAC TG ACACAGGAAG GGGCTGTGAA	120
10	CCCTAAGTGG GTGTATCTC CTCCRACCGA GGCTTCTMAC CCTGGAGATG GCAGTTACTC	180
	CTGGCCATGG TTGCTGAGCA TGGGCAGACC AGTGGAGGCC ACCCTACTGT GTTATCTGCG	240
	CCTTCRATGA AGTGAGACCC TTGGGGAGAA CGGGCTGTGG ATGAAGGAGT GGACTGCAGC	300
15	CTTGGCCTAG CCACTGGGCT GGGATCTTCT GGGTCATGTG ACTGTGTATC CAGGAGCAGA	360
	AACTTGTATT CTCAGGATTC AGGATCTACC CAGCACCAAA GATGTATTTT CAGGAGAACA	420
	GACCTAGAAA TGGGCCTGTC TGGCATTTCA GAGTCAGGCA AAGCAGGCAG GGCCAGGGAG	480
20	CTTCTGTGGG TCTACACAAG AAGGTTCTTG TGAGGGCTAT CAGTTGTTGC CTCTAGCTT	540
	GCTGGTAACT TTGGCGCCTC CGCCAAGCCC TGCCAGACTC CCCTGGCTGT GATGGCATTC	600
25	TGTGCCATCC TGCCTTGTC CCAGCCTCTG CAGGATGCCC TCCCTACCCA MCTYTYCCTG	660
	GGCCTTCCTT GTCCACTGGG CTGGATTCTT GTTCAAACCA CTGGACTGGC AGGGCAACGA	720
	CTTCTTCCCA CCTCAAGATG AGGTCCCTCG CCCCTTGTCT TGGCATAAAA ACACCTTTAA	780
30	AGCATGAGCC ATGTGCTTCT TTGCCCTTCT CTGTCTGTGT CCAATCTTCT GCCTCCCAGT	840
	CACTCCCTGG GGA CTATGGG ATCACTGTCC CCCACCTGT GTGGCCACAC CATGTGTCCT	900
35	GTCAATCCAG AACTGCCTCT GAGCTCCAGG CTGACCACAG ATCAGCCACA GCCTGATGCC	960
	TGCAGCCCCA CTTTGCTCAC CCTTCCCCTC CCTCCTCCT TCCTTCCACA CAGCAAGCCT	1020
	ACCTTITYTC ATCCATGCTC ACCATAGCCC CCTTCCTTGT GACCTGGACC CTCCATGTGA	1080
40	CCTGGCTGAG ACTGTCAGCC TCCTGGAGGA GTGGGTCCA CCTTCTTCTT GCCCTATGCA	1140
	GTGCAAGCTT CACTTCTCAC CCAGCAAGST TGA CTCTCT GCCTCCATGT CTCTGGGGCT	1200
45	TTGCTGTGTC CCTGAAACCT AGCTGGGCTG GTCTTGCTCC CAGCTTGCTT CCCCCTCCTC	1260
	GGATGTCCCT TTGCAGGCCC CTGTCTTCC TCCGGCACCA GTGTCTTGG CTGCCATGGC	1320
	AAGCTCATCA GGGGCTTGTA CCTGGTTCAC CAAGCATGGT AGCAGCTGCC TGCATTGTAT	1380
50	CTCCATCTGG TCACTGCAGG TGCCAACCCT TCATCCCCCA TGTTTTCTTG GGCCATGGAG	1440
	GGCTGACCTC CGTTTCTGGG GAATGTGGCT GAGCTGTGGT AACCAGCTAC ACCCCAGGTG	1500
55	CTCTTTCCAT GGTGGTGCTT GCTCATCTTG CTGATGCAAA CTAGGAAGTT AGGCTGCATC	1560
	TGGAGTGGC TTTCGCTGGA GAGGTGCTTT GCTGTCTCTC AGACTCAGTC ACTGTGTGCC	1620
60	CTCCCCGCTT CTCTTATCTC CATGGCTGTT TGCAGCTCTC CCAGGTACTT TGGGGTCTGA	1680

GCTGGAATTC CTTTGTGGTT TGCTCTCTG CTCTCTACTC TTGTATTAAG AAGGATTCCA 1740
CAAAGGGAGA GTGGCATCCC TGCTGCTGCT GTGCCAGACC AGAGTTTCCT GAGGGGCCCT 1800
5 GACCCTAACC CTCCAGCTCA GCCCTGTACA CCTGACCCTG TAAATGAGTG GGGTTTGCTG 1860
ACTGTAATCC CTGACACCAG TAAAACCAA AGGACTCTTG GGGGCTCAGT GTGAGAGCCA 1920
GGGTTACCTA CTCTGCCAAG TGAGGACAAA CTGCTAGGCT GTATCCATA ATTTTCAGGAT 1980
10 GAGAAACATT AACATAAAA ATTTGTAGTA AACATAACCT CATGANGACT AAAAAAAAAA 2040
AAAAACTYGG GGGGGGCCG GTAACCCATT GGGCCCTTG GGGGGNGTT TTAAAATT 2098

15

(2) INFORMATION FOR SEQ ID NO: 149:

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1847 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 149:

TCGACCCACG CGTCCGAAC T GAGCGGCGG CGGAGCCGG TTGGKGTCTG GTCTTCGGGT 60
30 CGGCCCCGCG GACCAGACG TGCCCCCGG GCGGGGAGAA GATGGTGCK AGCGGCCTCG 120
GGCCCGCCAC GCGCCGCCAC GAGTGAGCCC AGCGCGACG CGGCGCTCCG CCGAGCAGCT 180
GGCCCGGCTG GCGCCGGGC GCGCANTGCC CGCCGGGGCG GGGTGGAGCT GATCAGAATA 240
35 ATGTTTCAGCA TCAACCCCT GGAGAACCTG AAGGTGTACA TCAGCAGTCG GCCTCCCCTG 300
GTGGTCTTCA TGATCAGCGT AANGCCCATG GCCATAGCTT TCCTGACCCT GGGCTACTTC 360
40 TTCAAAATCA AGGAGATTAA ATCCCCAGAA ATGGCAGAGG ATTGGAATAC TTTTCTGCTA 420
CGGTTCAATG ATTTGACTT GTGTGTATCA GAGAATGAAA CCTCAAGCA TCTCACAAAC 480
GACACCACAA CTCCGAAAG TACAATGACC AGCGGGCAGG CCCGAGCTTC CACCCAGTCC 540
45 CCCCAGGCC TGGAGGACTC GGGCCCGTG AATATCTCAG TCTCAATCAC CCTAACCCCTG 600
GACCCACTGA AACCTTCGG AGGTATTCC CGCAACGTCA CCCATCTGTA CTCAACCATC 660
50 TTAGGGCATC AGATTGGACT TTCAGGCAGG GAAGCCACG AGGAGATAAA CATCACCTTC 720
ACCCCTGCTA CAGCGTGGAG CTCAGATGAC TGCGCCCTCC ACGGTCACTG TGAGCAGGTG 780
GTATTCACAG CTGTCATGAC CTCACGGCC AGCCCTGGGG TGTTCCTCGT CACTGTACAG 840
55 CCACCCACT GTGTTCTGA CAGTACAGC AACCCACGC TCTGTACAA GATCTTCACA 900
ACTGCCAGAG ATGCCAACAC AAAATACGCC CAAGATTACA ATCCITTCCTG GTGTTATAAG 960
60 GGGGCCATG GAAAAGTCTA TCATGCTTAA AATCCCAAGC TTACAGTGAT TGTTCAGAT 1020

GATGACCGTT CATTAATAAA TTGCACTCTC ATGCACACCA GTTACTTCCT CTTTGTGATG 1080
GTGATAACAA TGTTTTGCTA TGCTGTTATC AAGGGCAGAC CTAGCAAATT GCGTCAGAGC 1140
5 AATCCTGAAT TTGTGCCGA GAAGGTGGCT TTGGCTGAAG CCTAATTCCA CAGCTCCTTG 1200
TTTTTTGAGA GAGACTGAGA GAACCATAAT CCTTGCCTGC TGAACCCAGC CTGGGCCTGG 1260
10 ATGCTCTGTG AATACATTAT CTTCGATGT TGGGTTATTC CAGCCAAAGA CATTTCAAGT 1320
GCCTGTAAC TATTGTACA TATTTATAAA AATCTATTCA GAAATGGTC CAATAATGCA 1380
CGTGCTTTGC CCTGGGTACA GCCAGAGCCC TTCAACCCCA CCTTGGACTT GAGGACCTAC 1440
15 CTGATGGGAC GTTCCACGT GTCTCTAGAG AAGGATTCTT GGATCTAGCT GGTCACGACG 1500
ATGTTTTCAC CAAGGTCACA GGAGCATGTC GTCGCTGATG GGGTTGAAGT TTGGTTTGGT 1560
20 TCTTGTTC ACGCAATATG TAGAGAACAT TTGAAACAGT CTGCACCTTT GATACGGTAT 1620
TGCATTCCA AAGCCACCA TCCATTTTGT GGATTTTATG TGTCTGTGGC TTAATAATCA 1680
TAGTAACAAC AATAATACCT TTTCTCCAT TTGCTTGCA GGAAACATAC CTTAAGTTT 1740
25 TTTTGTTTT TTTTGTTTT TTTGTTTT GTTTCTTT ATGAAGAAA AATAAAATAG 1800
TCACATTTA ATACTACCA AAAATGGACA AAAAAAGTCG AGGGGGG 1847

30

(2) INFORMATION FOR SEQ ID NO: 150:

35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1569 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 150:

GACGCTGACG AGAGAAGGCC TCTTCCTTGA GGGTTGGTGC TGTGTTGCAG TGACCGTGGC 60
45 GGATTACGCC AACTCGGATC CGGCGGTCGT GAGGTCTGGA CGAGTCAAGA AAGCCGTAGC 120
CAACGCTGTT CAGCAGGAAG TAAAATCTCT TTGTGGCTTG GAAGCCTCTC AGGTTCTCTG 180
AGAGGAAGCT CTTTCTGGGG CTGGTGAGCC CTGTGACATC ATCGACAGCA GTGATGAGAT 240
50 GGATGCCCAG GAGGAAAGCA TCCATGAGAG AACTGTCTCC AGAAAAAGA AAAGCAAGAG 300
ACACAAAGAA GAACTGGACG GGGCTGGAGG AGAAGAGTAT CCCATGGATA TTTGGCTATT 360
55 GCTGGCCTCC TATATCCGTC CTGAGGACAT TGTGAATTTT TCCCTGATTT GTAAGAATGC 420
CTGGACTGTC ACTTGCAC TGCCCTTTTG GACCAGGTTG TACCGAAGCA CTACACGCTG 480
GATGCTTCCC TGCCTTTGCG TCTGCGACCA GAGTCAATGG AGAAGCTGCG CTGTCTCCGG 540
60

	GCTTGTGTGA TCCGATCTCT GTACCATATG TATGAGCCAT TTGCTGCTCG AATCTCCAAG	600
	AATCCAGCCA TTCCAGAAAG CACCCCCAGC ACATTAAAGA ATTCCAAATG CTTACTTTTC	660
5	TGGTGCAGAA AGATTGTTGG GAACAGACAG GAACCAATGT GGAATTCAA CTTCAAGTTC	720
	AAAAACAGT CCCCTAGGTT AAAGAGCAAG TGTACAGGAG GATTGCAGCC TCCCGTTCAG	780
10	TACGAAGATG TTCATACCAA TCCAGACCAG GACTGCTGCC TACTGCAGGT CACCACCCTC	840
	AATTTTCATCT TTATTCCGAT TGTTCATGGGA ATGATATTTA CTCTGTTTAC TATCAATGTG	900
	AGCACGGACA TCGGGCATCA TCGAGTGAGA CTGGTGTTC AAGATTCCCC TGTCCATGGT	960
15	GGTCGGAAAC TGGCAGTGA ACAGGGTGTG CAAGTCATCC TGGACCCAGT GCACAGCGTT	1020
	CGGCTCTTTG ACTGGTGGCA TCCTCAGTAC CCATTCTCCC TGAGAGCGTA GTTACTGCTT	1080
20	CCCATCCCTT GGGGGCAGCC TCGAGTGTAG TCCATTAGTA ATCAGATTCC AGTTTGACA	1140
	GGGTGGCTGG ATGTATATC TCGTTAGTAA TGTACATGCT CTTCAGGTTC TAGGGCTCCT	1200
	GTTAGGGGAG GGAGAAATGT TGAATCAAGA GGGAAAACAA CTACTATGAT TTATAAACAT	1260
25	ATTTTAATGT AAAAATTGTC ATTTAAAAGG AGTGGCCCTG TTTTCTGTGT TAAAACCCCA	1320
	TTTGGTGCTA TTGAGTTTGT TCTTTATCTT TTTATCCCAG TGAAAATTGT TGATCTTGCT	1380
30	GTAGGGAAAA ATTAACTCTT TTGAATCTCC AAACAAGGAA GTTTCAGCAT TCCCTTATGG	1440
	ATCAGAGGAA CCTTAGAGGC CTGAAATGTG TGCTTCCAGT TTAGCTGCCC CTCAAATTCA	1500
	AGTGAATATT TTCCCTTCTC CCTTTACCCT TCTCCAGAAA TAAAGCAGGT GACAGGGTTT	1560
35	CAGAATCTT	1569

40 (2) INFORMATION FOR SEQ ID NO: 151:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 1540 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 151:

50	CCCACGCGTC CGGAAGGATT GACCAGTTAA CCAACATCTT AGCCCCCATG GCTGTTGGCC	60
	AGATTATGAC ATTTGGCTCC CCAGTCATCG GCTGTGGCTT TATTTCGGGA TGGAAGTTGG	120
55	TATCCATGTG CGTGGAGTAC GTCCTGCTCT GGAAGGTTTA CCAGAAAACC CCAGCTCTAG	180
	CTGTGAAAGC TGGTCTTAAA GAAGAGGAAA CTGAATTGAA ACAGCTGAAT TTACACAAAG	240
	ATACTGAGCC AAAACCCCTG GAGGGAATC ATCTAATGGG TGTGAAAGAC TCTAACATCC	300
60	ATGAGCTTGA ACATGAGCAA GAGCCTACTT GTGCCTCCA GATGGCTGAG CCCTTCCGTA	360

5 CCTTCCGAGA TGGATGGGTC TCCTACTACA ACCAGCCTGT GTTTCCTGGCT GGCATGGGTC 420
TTGCTTTCCT TTATATGACT GTCCTGGGCT TTGACTGCAT CACCACAGGG TACGCCTACA 480
CTCAGGGACT GAGTGGGTTT CATCCTCAGT ATTTTGATGG GAGCATCAGC TATAACTGGA 540
ATAATGGGAA CTGTAGCTTT TACTTGGCTA CGTCGAAAAT GTGGTTTGGT TCGGCAGGTC 600
10 TGATCTCAGG ATTGGCACAG CTTTCTGTG TGATCTTGTG TGTGATCTCT GTATTCATGC 660
CTGGAAGCCC CCTGGACTTG TCCGTTTCTC CTTTGAAGA TATCCGATCA AGGTTTCATT 720
AAGGAGAGTC AATTACACCT ACCAAGATAC CTGAAATTAC AACTGAAATA TACATGTCTA 780
15 ATGGGTCTAA TTCTGCTAAT ATTGTCCCGG AGACAAGTCC TGAATCTGTG CCCATAATCT 840
CTGTCAGTCT GCTGTTTGCA GCGTCATTG CTGCTAGAAT CGGTCTTGG TCCTTTGATT 900
20 TAACTGTGAC ACAGTTGCTG CAAGAAAATG TAATTGAATC TGAAAGAGGC ATTATAAATG 960
GTGTACAGAA CTCCATGAAC TATCTTCTTG ATCTTCTGCA TTTTCATCATG GTCATCCTGG 1020
CTCCAAATCC TGAAGCTTTT GGCTTGCTCG TATTGATTTC AGTCTCCTTT GTGGCAATGG 1080
25 GCCACATTAT GTATTTCCGA TTTGCCCAA ATACTCTGGG AAACAAGCTC TTTGCTTGCG 1140
GTCCTGATGC AAAAGAAGTT AGGAAGGAAA ATCAAGCAA TACATCTGTT GTTTGAGACA 1200
30 GTTTAACTGT TGCTATCCTG TTAGTAGATT ATATAGAGCA CATGTGCTTA TTTTGTACTG 1260
CAGAATTCCA ATAAATGGCT GGTGTTTGT CTCTGTTTIT ACCACAGCTG TGCCTTGAGA 1320
ACTAAAAGCT GTTTAGGAAA CCTAAGTCAG CAGAAATTAA CTGGATTAA TTTCCCTTATG 1380
35 TTGAGGGCCA TGGRAAAAA ATTGGGAAA GGAAAACTC AGTTTAAAT ACGGGAGACT 1440
ATAATGGATA AACTGTRATT CCCCTATTTC TCATGAGTAG ATACAATCTT ACGTAAAAGA 1500
40 GTGGTTAGTC ACGTGAATTC AGTTATCATT TGACAGATTC 1540

45 (2) INFORMATION FOR SEQ ID NO: 152:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1719 base pairs
(B) TYPE: nucleic acid
50 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 152:

55 TACTTATGAG GTCAATTGGA AATAAGAACA CCAATTTACT GGGTCTAGGA TTTCAAATAT 60
TACAGTTGGC ATGGTATGGC TTTGGTTTCA AACCTTGGAT GATGTGGGCT GCTGGGGCAG 120
TAGCAGCCAT GTCTAGCATC ACCTTTCTG CTGTCAGTGC ACTTGTTTCA CGAACTGCTG 180
60

	ATGCTGATCA ACAGGGTGTC GTTCAAGGAA TGATAACAGG AATTTCGAGGA TTATGCAATG	240
	GTCTGGGACC GGCCCTCTAT GGATTCAATT TCTACATATT CCATGTGGAA CTTAAAGAAC	300
5	TGCCAATAAC AGGAACAGAC TTGGGAACAA ACACAAGCCC TCAGCACCAC TTTGAACAGA	360
	ATTCCATCAT CCCTGGCCCT CCCTTCCTAT TTGGAGCCTG TTCAGTACTG CTGGCTCTGC	420
10	TTGTTCGCTT GTTTATTCCG GAACATACCA ATTTAAGCTT AAGGTCCAGC AGTTGGAGAA	480
	AGCACTGTGG CAGTCACAGC CATCCTCATA ATACACAAGC GCCAGGAGAG GCCAAAGAAC	540
	CTTTACTCCA GGACACAAAT GTGTGACGAC TGAAATCAGG AAGATTTTTC TATCAGCACC	600
15	CAGGTCTTAG TTTTCACCTC TAGTTCCTGA TGTACATTCC ATTTCCATCC ACAGTGTACT	660
	TTAAGATTGT CTTAAGAAAT GTATCTGCAT GAACTCCGTG GGAAGTAAAG GAAGTGGGAA	720
20	CTTAGAACCA GACAGTTTTC CAAAGATGTT ACAATTTCTT TTGAAAAACC TTTTGTATTAT	780
	TAGCACCAAT TTCTYGCCAC TAAGCTATTT GTTTTATTAT ACATCCTTTA ATTAAAACT	840
	ATATATGTAA CTTCTTAGAT ATTAGCAAAT GTCTCTGCTA CCATTTCTTT AAGGTGTTGA	900
25	GCTTTAACTC TATGCTGACT CAGTGAGACA CAGTAGGTAG TATGGTTGTG GACCTATTTG	960
	TTTTAACATT GTAAAATTTT GAGTCAGATT TTAATATTGT AAAATCTTGG GTCAAATAAT	1020
30	TCAAAGCCTT AATGCAGATG CACTAAAACA AAGAAATGGT AAATGAATTG TTTGCATTTA	1080
	AAAAAAAAAA CTCCTAAGAA AACTGTACTA AATCTGAATC ATGTTTGTAG CTTGTTTGCA	1140
	GTACTTTTAA ACATTATTCA CTACTGTTTT TGAAGTGAGA AAGTATCAGC CATTTAGCAT	1200
35	TTAAGTTGGG GTATTTAGAG CCTGTAATCT AAATGCTGGC TCAAATTTAT TCCCCAGCTA	1260
	CTTCTTATAC CACTATTCTT TTAATGTTTG CATAATCATA AGCACCTCAA CACTTGAATA	1320
40	CATAATCTAA AAATTATATA GTAAAGCTGG TAGCCTTGAA AATGTCAGTG TGATATCTAT	1380
	TATGTAGATA AATATATATA GTGGCCTTTC AGGACTGTCA CAGTAACACT TTATTTACAG	1440
	AGCTAATGTT TGTCTAAAT TTTCAAGACC CTAGAGGAGA GCTTTATACA ATTACCGATG	1500
45	TGAATTTCTC TAAAGTGAT ATTTTGTGT CCAGTTATAT TATTTAAAAA AGTGTACTTT	1560
	TGTAAAAATT GTATATAAAG AACTGTATAG TTTACACTGT TTTCACTCTG TGTGTGGTTA	1620
50	TTGCTTAATG CTTTTTAAAC TTGGAACACT CACTATGGTT AAATAAGGTC TTAAAAGAAA	1680
	TGTAAATATT YGTPTAATAA AGTTAAATAT TTTAATGAT	1719

55

(2) INFORMATION FOR SEQ ID NO: 153:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 863 base pairs

(B) TYPE: nucleic acid

60

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: 'SEQ ID NO: 153:

5
GGCACGAGGG AAGCCGGGAC GATGTCCGCA TGACAACCGA CGTTGGAGTT TGGAGGTGCT 60
TGCCTTAGAG CAAGGGAAAC AGCTCTCATT CAAAGGAACT AGAAGCCTCT CCCTCAGTGG 120
10 TAGGGAGACA GCCAGGAGCG GTTTTCTGGG AACTGTGGGA TGTGCCCTTG GGGGCCCGAG 180
AAAACAGAAG GAAGATGCTC CAGACCAGTA ACTACAGCCT GGTGCTCTCT CTGCAGTTCC 240
TGCTGCTGTC CTATGACCTC TTTGTCAATT CCTTCTCAGA ACTGCTCCAA AAGACTCCTG 300
15 TCATCCAGCT TGTGCTCTTC ATCATCCAGG ATATTGCAGT CCTCTTCAAC ATCATCATCA 360
TTTTCTCAT GTTCTCAAC ACCTTCGTCT TCCAGGCTGG CCTGGTCAAC CTCCTATTCC 420
20 ATAAGTTCAA AGGGACCATC ATCCTGACAG CTGTGTACTT TGCCCTCAGC ATCTCCCTTC 480
ATGTCTGGGT CATGAACTTA CGCTGGAAAA ACTCCAACAG CTTCATATGG ACAGATGGAC 540
TTCAAATGCT GTTTGTATT CAGAGACTAG CAGCAGTGT GTACTGCTAC TTCTATAAAC 600
25 GGACAGCCGT AAGACTAGGC GATCCTCACT TCTACCAGGA CTCCTTGTGG CTGCGCAAGG 660
AGTTCATGCA AGTTCGAAGG TGACCTCTTG TCACACTGAT GGATACTTTT CCTTCTGGA 720
30 TAGRAGGCCA CATTTGCTGC TTTGCAGGGG AGAGTTGGGC CCTATGCATG GGGCAAACA 780
GGTGGGATTT TCCAAGGAA GGGTTCAGAA TTAGGCNIGT TGTTTCAGCC ATTTCCAAGG 840
35 AAGGGGAAGG GTTTCCTTNC CCT 863

(2) INFORMATION FOR SEQ ID NO: 154:

40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1101 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 154:

50 AACAGCAAAA AAGAATGATT TCTTCTGAAA TTGTGGAACA TGAGGATTCA AGTTTTTATT 60
TTGTTACTAG GTGCTGGAGG AACATCCCAG TTCACAAAGC CCCCATCTCT TCCTCTGGAG 120
CCAGAGCCTG CGGTGGAATC AAGTCCAAC GAAACATCAG AACAAATAAG AGAGAAATAA 180
55 GAATAGAATG AATGACCCCA AAATARGGTT TTCTTGGGCG AGGATGTGCT GGATTAGGAA 240
AGGTGACATG ACACAGGCAG AGCAGAGTGG CACCCACCAC AGAATACAGT GTGTGTTATT 300
ACGAGGAGCC AGCAGTTGAG CCTAAGGTCC TTCTACCTAC CTGGTATTGG CATTTGAGGT 360
60

CGGAAACCCCT CTACTGCCCC ATAAGCCAGG AAAAGTGAAA AGAGAACACA GTTCCTTTAA 420
 GAACTGGCAG CAAGGCTTGA GGCCTTATGT ATGTAGCTGA GTCAGCAAGG TACATGATGC 480
 5 TGTCTGCTTT CAAAAGGACT TTTCTCTCCT AGCTGACTGA CTCCTTCCTT AGTTCAAGGA 540
 ACAGCTGAGA CAGACCTCTG CTGAGTAGCT CTGTGATGAC AAAGCCTTGG TTAACTGAG 600
 GTGATCCTCA GGTGTGAGG TTTATTAGTC CCCAAGGCAA ACACAAATAT TAGATTAATA 660
 10 ATCCAACCTT AATAGTATAC ATTTAAAAGA AAAAAACAA AAGCCCTGGA AGNTTGAGGC 720
 CAAGCCTGCT GAGTATGCA GCTGCATTTG CCCAAAGGGA ATCCAGAACA AGTCCCTCCC 780
 15 TGTATTTTGT TCTTGAGAGG GGTGAGTCTA GAAGCTAGAT CCTATCAGGA TGAGGAGCAG 840
 CAGCCCAGGG CTGTCTGGA TCAGCACCAA CGATTTTAAA GAAAAAGGA AGAGTTTCTT 900
 AGATGAGTAA TTGTTATTGA AGATAGTCAG TGATAACCAC TGACCAGATG CTATCAATAC 960
 20 ACTATGTGTC CTMTTGTAGAA TAAAGATTAC ATATCATCAT TCCTTTGGGG AAAATGTGA 1020
 TTCAGGTATA AAAACAAGAG ATTATAATAA AAAANTAAAA GAACCCTAAA AAAAAAAAC 1080
 25 CTCGTGCCGA ATCCCTGCA G 1101

30 (2) INFORMATION FOR SEQ ID NO: 155:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2031 base pairs
 (B) TYPE: nucleic acid
 35 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 155:

40 CAATTAAACC GTTTGAGGCC TAGGTTGTTT GGCAAGCCCC NGGCCTAAAG TTTTAATTGC 60
 GCAGAGCCAA GGCCTGAAA GGAAGGAAA GGGGAGGTA GCGGAGGGT AGCAGGTGAG 120
 TTCCTAGGGC TGAAGGTTT AGCAGCAGCC TGGTGAGTG CCCTGTCATC AAGACAAACC 180
 45 CACGGTCTC CTGGTGCCT ACCAAGCTTG GTTGTACAA AAGCAAGGTG GGAGTCTATT 240
 TTTGTACATG AGATACATCA CACTTACCTG TGGCCAGTA TTGTGAAGTG AGTCTGAGTT 300
 50 GTTTACTACTG ATGCCCTCCC TGCCACCAC AAATTGTGTA CATAGTCTTC AGAATGATAC 360
 CACCCCTTTC CCCAGCTCCC AACCAAGAGC TGGTTCAGG CCTGTGTTAT ATGTCATATT 420
 TAGCGTTTTT ATATATGACC TTTGATTTCT GTTGTGTA TTTAGCACA GTGTATGCAC 480
 55 CTTCAATTAA ATACATCTGT GTGCATACAG ATACGCATAT ATGTGTGTC GTATGCATAT 540
 ATCTCTCATC TGTAGTTTCC AAGAGTTCAG CTGAAGCAGA TGGAGTCCTG CAGCCAGGA 600
 60 GACACCCTGC ATCCCTGCTA ATAGTGTGTA CCACAAGTAT TAGTGAGTCT TCCTTATTAA 660

TATTTTCATT TCAGAAGACT GAAGCAAAGC TGATAGTGTT TGCTGTTTCT TTGGCAGCTA 720
AGTGAGGGTC TTGGGATGAC TTGCTGTGTT CCTCAAGCTG CACTTTGGGG CCATCTCTGC 780
5 AGTATTAAGC CCCCTTTTGG CTGGGTGGTA CTCTGTCTGT GCCTGTGTGT GTGTGTGATA 840
GTCACTCTTG CATGGCTTCC ATGTCTGGTT TGTGGCATTT GGGGATAAGT GCTGAACCAG 900
10 AGCATTTCGA GTTGTGTTGA GGCCTCGTTG CCAATGATAG ATCACTCCTG TTGACCTGGT 960
ATGTCTGCTT GCTGTCTGCT TTTCTTTGCT TTCTCTTGGA AGAGGAAAGG ACTCTGGTCA 1020
GGCCAGGCT GAGTGAGATG AGCTGCAGCT GGCTCATGGC CTCTCTAGAG CAGAGAGAGG 1080
15 AGTATGTCAT TTTACTAAGT TCCTAAACAA ACATTTATGC AGGCAACACT CCTTGCAGAT 1140
CCAGAACTG AGGCACAATA GGGTTATGAC TTGCTCAAGA ATATGTAGCT GCTAGGGGGT 1200
20 AAATCAAGGC ATCACAATTT CTGTTTCAGCG GGCAGGAATA GGCTGTGAAT TGCTAGCACT 1260
TTTTTTTAA GCAATTACTT TTTGACTTGT TCCTCTGAAA GTGCAAGAGG CGTACACCTT 1320
TCCCAAATGT AGACTAGAAT CTGCAGGATG CCACCCACTG TATAGTTCTG CTTTCCAGA 1380
25 GAGGAAGAAC TTTTAGAAAC CAAATGATCT TAATGTGTAT TGCCACCCCC TGGCTTTTCC 1440
GGGTAGAAAA TTCACAGTAG GAATGATTGT TAAGAGAGAG TGCTTGGAAC CATGGGTTAA 1500
30 CAGGAAAGGC TACCTAACTT CACATATCTG CAACCAGAGC AGCCACCAAG CATTACTTAG 1560
CAGCAGGAAA ATGATTGTAT TTGAGTTCTT GTGTGTCCAA AACTGAGGCA CCATGTTCTT 1620
TGAAAACATG CCACCTCAAG GCTGGGCGCG GTGGCTCACA CCTGTTAATC CCAGCACTTT 1680
35 GGGAGGCCGA GGGGGCGGA TCACCGGAGT CGGGGAGTTT GAGACCAGCC TGGACCAACA 1740
TGGGAGAAAC CCCATCTCTA CCTAAAAATA CAAAATTAGC CGGGCGTGGT GGCATGCGCC 1800
40 TATAATCTCA GCTACTTGGG AGGGYTGAGG CAGGRGAATT GCTTGAACCC RGGANGGCGG 1860
AGGTTTCCGG TTGAGTTGAG GATCGTGCCA TTGCACTTCC GGGCCTTGGG GCAACAACAG 1920
CAAAAAYTCC GTCTTCAAMW MRTGCCGAAT TCGATATCAA GCTTATCGAT ACCGTCGACC 1980
45 TCGAGGGGGG GCCCGGTACC CAATTCGCCC TATAGNGATC GTATTACAAT C 2031

50

(2) INFORMATION FOR SEQ ID NO: 156:

(i) SEQUENCE CHARACTERISTICS:

55

- (A) LENGTH: 1981 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 156:

	CCTGCACCCCT GAGCCCTTCA CCCCTCCGAG TTCCCCCAG GTTGGCTTCC TTCGATTCCCT	60
	TTTCTTGGTA TCAACGTTTG ATTGGAAGAA CAACCCCTC TTTGTCAACC TCAATAATGA	120
5	GCTCACTGTG GAGGAGCAGC TCGGGCACAG CTCMCCGTYA TGGTCATTGT TACCCCCCAA	180
	GACCGCAAAA ACTCTGTGTG GACACAGGAT GGACCCTCAG CCCAGATCCT GCAGCAGCTT	240
10	GTGGTCTCGG CAGCTGAAGC CCTGCCCATG TTAGAGAAGC AGCTCATGGA TCCCCGGGA	300
	CCTGGGGACA TCAGGACAGT GTTCCGGCCG CCCTTGGACA TTTACGACGT GCTGATTCCG	360
	CTGTYTCTC GCCATATCCC GCGGCACGC AGGCTTGTGG ACTCGCCAGY TGCTCTCTC	420
15	TGCCGGGGCC TGCTCAGCCA GCCGGGGCCC TCATCCCTGA TGCCCGTGCT GGGTNATGAT	480
	CCTNCTCAGC TCTATCTGAC GCAGCTCAGG GAGGCCTTTG GGGATCTGGC CCTTTTCTTC	540
20	TATGACCAGC ATGGTGGAGA GGTGATTGGT GTCTCTGGA AGCCACCAG CTTCCAGCCG	600
	CAGCCCTTCA AGGCCTCCAG CACAAAGGGG CGCATGGTGA TGTCTCGAGG TGGGGAGCTA	660
	GTAATGGTGC CCAATGTTGA AGCAATCCTG GAGGACTTTG CTGTGCTGGG TGAAGGCCTG	720
25	GTGCAGACTG TGGAGGCCCC AAGTGAGAGG TGGACTGTGT GATCCAGCT CTGGAGCAAG	780
	CTGTAGACGG ACAGCAGGAC ATTGGACCTC TAGAGCAAGA TGTCAGTAGG ATGACCTCCA	840
30	CCCTCCTTGG ACATGAATCC TCCATGGAGG GCCTGCTGGC TGAACATGCT GAATCATCTC	900
	CAACAAAACC CAGCCCCAAC TTTCTCTCTG ATGCTCCAGC ATTGGGGCAG GGGCATGGTG	960
	GCCCATGTAG TCTCCTGGGC CTCACCATCC CAGAAGAGGA GTGGGAGCCA GCTCAGAGAA	1020
35	GGAATGAAC CCAGGAGATC CATCCACCTA TTAGCCCTGG GCCTGGACCT CCCTGCGATT	1080
	TCCCACTCCT TTCTTAGTCT TCTTCCAGAA ACAGAGAAGG GGATGTGTGC CTGGGAGAGG	1140
40	CTCTGTCTCC TTCTGTCTGC CAGGACCTGT GCCTAGACTT AGCATGCCCT TCACTGCAGT	1200
	GTCAGGCCTT TAGATGGGAC CCAGCGAAAA TGTGGCCCTT CTGAGTCACA TCACCGACAC	1260
	TGAGCAGTGG AAAGGGGCTA TATGTGTATG AATAGACCAC ATTGAAGGAG CACAATGCCC	1320
45	TCCTGTGTTG ATGCCACTTC CCAGGGTGA GACAGTGGAA AAGAACCGAG GACAGGAAAG	1380
	GATTGGGTAG GTGAAGGGT CAGGGGACTG GTAGTCACCC AATCTTGGAG AGGTGCAAAA	1440
50	AGCACTGGGG GCTACCCGTT AGCTGCATCT GCCCTGGCTG TTTGCCCGTT CATGTCACAA	1500
	ACTGCCACTA CTATGTACCT GCAGTGGGT TGCAGAGATG GGGGAGACTC AAGTCTTACT	1560
	CCCCAGGAGC TCCCAGGGCC CAAGGAGGAG AATGCTGCCT CCTTTCAGTC TGGTCTACAC	1620
55	CCACTTTCCTG GTAGCCTCTC TGCTTCCTGT AATTCCTGGCT GTTTTTCAG ACTCAGCTCA	1680
	AATAGTCCCC CTCCTTAAGC CCATCCCTCG CCCCAGCCT GAGGTGATCT TTCCCTCCTC	1740
60	TGAACTATTA GAGCAGTTAC TGTCTGTTC AATCGTTTGG CAGGCACACA CAGTGGCATA	1800

AATTCCTATTG TTTTGAAGTC TGATTTAAAA TTAAATTGCA GCTGGGCGTG GTGGCTCATG 1860
CTTGTAATCC CAACACTTAG GGAGTMAGGR GAATCACTTG ASCYCAGGAG TYCTAGACCA 1920
5 ATCTGGGCAA MAGAGAGACC CCATCTCTTT TAAATAAAAA GTTAAATTGC TTAACAAAAA 1980
A 1981

10

(2) INFORMATION FOR SEQ ID NO: 157:

(i) SEQUENCE CHARACTERISTICS:
15 (A) LENGTH: 915 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 157:

GAATTCGGCA CGAGCGCGGC CATGGCGCTC CTGCTTTCGG TGCTGCGTGT ACTGCTGGGC 60
GGCTTCTTCG CGCTCGTGGG GTTGGCCAAG CTCTCGGAGG AGATCTCGGC TCCAGTTTCG 120
25 GAGCGGATGA ATGCCCTGTT CGTGCACTTT GCTGAGGTGT TCCCGCTGAA GGTATTTGGC 180
TACCAGCCAG ATCCCCTGAA CTACCAAATA GCTGTGGGCT TTCTGGAACT GCTGGCTGGG 240
30 TTGCTGCTGG TCATGGGCCC ACCGATGCTG CAAGAGATCA GTAACCTGTT CTTGATTCTG 300
CTCATGATGG GGGCTATCTT CACCTTGGCA GCTCTGAAAG AGTCACTAAG CACCTGTATC 360
CCAGCCATTG TCTGCCCTGG GTTCCCTGCTG CTGCTGAATG TCGGCCAGCT CTTAGCCCAG 420
35 ACTAAGAAGG TGGTCAGACC CACTAGGAAG AAGACTCTAA GTACATTCAA GGAATCCTGG 480
AAGTAGAGCA TCTCTGTCTC TTTATGCCAT GCAGCTGTCA CAGCAGGAAC ATGGTAGAAC 540
40 ACAGAGTCTA TCATCTTGTT ACCAGTATAA TATCCAGGT CAGCCAGTGT TGAAAGAGAC 600
ATTTTGTCTA CCTGGCACTG CTTTCTCTTT TTAGCTTTAC TACTCTTTTG TGAGGAGTAC 660
ATGTTATGCA TATTAAACATT CCTCATGTCA TATGAAAATA CAAAATAAGC AGAAAAGAAA 720
45 TTAAATCAA CCAAAATTCT GATGCCCAA ATAACCACTT TTAATGCCTT GGTGTAAGTA 780
TACCTCTGAA CTTTTTCTG TGCCTTTAAA CAGATATATA TTTTTTTTWA ATGAAAATAA 840
50 AACCATATAT CCTATTTTAT TTCCTCCTTT TAAACCTTA TAACTATAA MAAAAAAA 900
AAAAAAA CTCGA 915

55

(2) INFORMATION FOR SEQ ID NO: 158:

(i) SEQUENCE CHARACTERISTICS:
60 (A) LENGTH: 2117 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 158:

	AGAGCGAAGC GAGGGTGGCG CGGGTCCGG CATGAAGCTG GGCCGGGCCG TGCTGGGCCT	60
	GCTGCTGCTG GCGCCGTCCG TGGTGCAGGC GGTGGAGCCC ATCAGCCTGG GACTGGCCCT	120
10	GGCCGGCGTC CTCACCGGCT ACATCTACCC GCGTCTCTAC TGCTCTTCG CCGAGTGCTG	180
	CGGGCAGAAG CGGAGCCTTA GCCGGGAGGC ACTGCAGAAG GATCTGGACG ACAACCTCTT	240
15	TGGACAGCAT CTTGCAAAGA AAATCATCTT AAATGCCGTG TTTGGTTTCA TAAACAACCC	300
	AAAGCCCAAG AAACCTCTCA CGCTCTCCCT GCACGGGTGG ACAGGCACCG GCAAAAATTT	360
	CGTCAGCAAG ATCATCGCAG AGAATATTTA CGAGGGTGGT CTGAACAGTG ACTATGTCCA	420
20	CCTGTTTG TG GCCACATTGC ACTTTCCACA TGCTTCAAAC ATCACCTTGT ACAAGGATCA	480
	GTTACAGTTG TGGATTGAG GCAACGTGAG TGCTGTGCG AGGTCCATCT TCATATTTGA	540
25	TGAAATGGAT AAGATGCATG CAGGCCTCAT AGATGCCATC AAGCCTTTCC TCGACTATTA	600
	TGACCTGGTG GATGGGTCT CCTACCAGAA AGCCATGTTT ATATTCTCA GCAATGCTGG	660
	AGCAGAAAGG ATCAGAGATG TGGCTTTGGA TTTCTGGAGG AGTGGAAAGC AGAGGAAGA	720
30	CATCAAGCTC AAAGACATG AACACGCGTT GTCTGTGTCG GTTTTCAATA ACAAGAACAG	780
	TGGCTTCTGG CACAGCAGCT TAATTGACCG GAACCTCATT GATTATTTTG TTCCCTTCCT	840
35	CCCCCTGGAA TACAAACACC TAAAAATGTG TATCCGAGTG GAAATGCAGT CCCGAGGCTA	900
	TGAAATTGAT GAAGACATG TAAGCAGAGT GGCTGAGGAG ATGACATTTT TCCCCAAAGA	960
	GGAGAGAGTT TTCTCAGATA AAGGCTGCAA AACGGTGTTC ACCAAGTTAG ATTATTACTA	1020
40	CGATGATTGA CAGTCATGAT TGGCAGCCGG AGTCACTGCC TGGAGTTGGA AAAGAAACAA	1080
	CACTCAGTCC TTCCACACTT CCACCCCCAG CTCCTTTCCC TGGAGAGGGA ATCCAGTGAA	1140
45	TGTTCTCTGT TGATGTGACA GGAATCTCTC CTGGCATTGT TTCCACCCCC TGGTGCCTGC	1200
	AGGCCACCCA GGGACCACGG GCGAGGACGT GAAGCCTCCC GAACACGCAC AGAAGGAAGG	1260
	AGCCAGCTCC CAGCCCCTC ATCGCAGGGC TCATGATTTT TTACAAATTA TGTTTTAATT	1320
50	CCAAGTGTTT CTGTTTCAAG GAAGGATGAA TAAGTTTAT TGAAAATGTG GTAACCTTAT	1380
	TTAAAATGAT TTTTAACATT ATGAGAGACT GCTCAGATTC TAAGTTGTTG GCCTTGTTG	1440
55	TGTGTTTTTT TTTAAGTTCT CATCATTAT ACATAGACTG TGATGTATCT TTAAGTGAAA	1500
	TGAGCCCAAG CACACATGCA TGGCATTGTG TCCACAGGAG GGCATCCCTG GGGATGTGCC	1560
60	TGGAGCATGA GCCAGCTCTG TCCCAGGATG GTCCCAGCGG ATGCTGCCAG GGGCAKTGAA	1620

GTGTTTAGGT GAAGGACAAG TAGGTAAGAG GACGCCTTCA GGCACCACAG ATAAGCCTGA 1680
AACAGCCTCT CCAAGGGTTT TCACCTTAGC AACAAATGGA GCTGTGGGAG TGATTTTGCC 1740
5 CACACTGTCA ACATTTGTTA GAACCACTCT TTTGAAAGAA AAGTATTTCC AACTTGTTCAC 1800
TTGCCAGTCA CTCCTTTTGG CAAAAGGTGG CCCTTCACTG TCCATTCCAA ATAGCCACAA 1860
CGTGCTCTCT GCTGGATTCT AAATTATGTG AATTTTGCCA TATTAAATCT TCCTCATTTA 1920
10 TACTATTATT TGTTACGTTT AATCAGAATC CCCGAAACCT CCTATAAAGC TTAGCTGCCC 1980
CTTCTGAGGA TGCTGAGAAC GGTGTCTTTC TTTATAAATG CAAATGGCTA CCGTTTACAA 2040
15 ATAAAATTTT GCATGTGCAA AAAAAAAAAA ANAAAAAAAA AAAATCCCGG GGGGGGGCCG 2100
GTAACCAATT TGNCCCC 2117

20

(2) INFORMATION FOR SEQ ID NO: 159:

(i) SEQUENCE CHARACTERISTICS:
25 (A) LENGTH: 2395 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 159:

TGTTCTTAA TCCCTTTTCT AAAAAGGGGG GAAATCCGG ATGGATTTTA GGGATTGGTC 60
TGGTGTGAGC TGTGTTTAT TGCACACCTA AATCCTGATT ATAGGCTTTT CATTTCTCCG 120
35 CAAAGCCTTT ATTTTGCCAG TTAAGCCAAA TGTGTTTTC AGAAAGTTAG TTATTTTCTC 180
CTCTTCTTTT CCTTCTTTC CTCCTTTTTT CCCGTCTGAC CCCAAACGTT ATTGTCCAAA 240
40 CATGACTGGA CAGCAGCTTT TGTTTCTTGA CCCTGTAATA TGACAGTCTG CTAATATTGA 300
CAGAAGGTGC AGTTTMTGGG TTATAGTCGT GATTTTCGCT AATCAATCAT ATTAGCAGGA 360
AAAAAAKGA CTGTTTCTG TTGTAAGTGA GTCTTAAGAA AAAGTGGCCC ATAGTTTAGT 420
45 GGACAATTTT CAAAGGCTTT AGTACCACCT GTATTTCAAA ATGGGGGACC CAAACTCCCG 480
GAAGAAACAA GCTCTGAACA GACTACGTGC TCAGCTTAGA AAGAAAAAG AATCTCTAGC 540
50 TGACCAGTTT GACTTCAAGA TGTATATTGC CTTGTATTC AAGGAGAAGA AGAAAAAGTC 600
AGCACTTTT GAAGTGTCTG AGGTTATACC AGTCATGACA AATAATTATG AAGAAAAAT 660
CCTGAAAGGT GTGCGAGATT CCAGCTATTC CTGGAAGT TCCCTAGAGC TTTTACAGAA 720
55 GGATGTGGTA CAGCTCCATG CTCCTCGATA TCAGTCTATG AGAAGGGATG TAATTGGCTG 780
TACTCAGGAG ATGGATTCA TTCTTTGGCC TCGGAATGAT ATTGAAAAA TCGTCTGTCT 840
60 CCTGTTTCT AGGTGGAAG AATCTGATGA GCCTTTTAGG CCTGTTTCTG CAAATTTGAG 900

	TTTCATCATG GTGACTATGA AAAACAGTTT CTGCATGTAC TGAGCCGCAA GGACAAGACT	960
5	GGAATCGTTG TCAACAATCC TAACCAGTCA GTGTTTCTCT TCAITGACAG ACAGCACTTG	1020
	CAGACTCCAA AAAACAAAGC TACAATCTTC AAGTTATGCA GCATCTGCCT CTACCTGCCA	1080
	CAGGAACAGC TCACCCACTG GGGCAGTTGG CACCATAGAG GRTCACCTCC GTCCTTATAT	1140
10	GCCAGAGTAG AGTACTGACC AGCAAAATGG AGAAGATCAG AGAATGCAGC AGCAGTTTTF	1200
	TTTCTTGTTT TCTTACCACT TTATTCTTTC AGAGTTTAAA GAAAATGGAC TCATGCACAG	1260
15	AACACTATGC ATTTTGAAAC TTGTTTCATCC TGGATTTTTF TAAATCATTT TTATCTCAGA	1320
	ACTTAAACAA AAATTAGATG TCGTGCACGG ACTGTGTGAA AGAAGATGCT TTGCATATTT	1380
	GCTGCACTGC ATCAGTATCT TACTAAAAAT GTGAAATGAA AGGACTATTG TACACTGAAA	1440
20	TGCTTAAATG TATCTGAAAG CACAAGGTGA TACTCATTTT TATGGTCTTC CCATTTGTGC	1500
	TGGTTTTTGC CTCTTTGACA TCTGTCATCA GTATTTAGAG GGTGAGAAGT GAATGTAACA	1560
25	GGTATAAATA ACATTTTAA AAACAATAAC TTTGCTATAA TCACAGTTGT TCCAGAGCAC	1620
	TGTCAGATAC ATTCTAATGA CCAGAACTGG TTTAAAAAAA GAAAATACAA CCATGGGAAA	1680
	GAAATCTTAA ATGAAAAACG CATCTCATTG TAGGCATTTT TGCCTCATAT TTTACTGGGC	1740
30	CATGTTTGTT TCCTGGTACT CATGTATTTT TTTTTCCAG ATCTCTTTCC CCAAGTTGCT	1800
	ATTGTAAGAG TATTCTGCTG CGTGTGGATG CAGTTATACA CATTAAGCA GATCTGGAGT	1860
35	CTGAAGTAGC TATAAAGCAG CTATAAACA GAAATACATG CATAGCTGCA GAAACCATGA	1920
	TAGGTAGAGG ACTTTTCTTT TGGTTTTGTT TTGTTTTGTT TTGTTTTGTT TTTGGTTTTA	1980
	CAGAGAAGAG ATTTTTATTA CAAAGAAAAA AATTCCAGTG AATGTGCAG AAATGCTGGT	2040
40	TTTACACCA TCCTAAAGAA AAACCTTACA AGGGTGTTTT GGAGTAGAAA AAAGTTATA	2100
	AAGTTGGAAT CTTAAATTGT AAAATTAACC ATTGAGTGTG AAAGTTCTAA AAGCAGAACT	2160
45	CATTTTGTGC AATGAACATA AGGAAAGACT ACTGTATAGG TTTTTTTTTT TTCTCCTTTT	2220
	AAATGAAGAA AAGCTTTGCT TAAGGGTTGC ATACTTTTAT TGGAGTAAAT CTGAATGATC	2280
	CTACTCCTTT GGAGTAAAC TAGTGCTTAC CAGTTTCCAA TTGTATTTAG CTTCTGGTTG	2340
50	GAATTTGAAA AAAAAAGAAA AAAAGAAAAA GAAAACCTAA ATAAATAGG TGAAA	2395

55 (2) INFORMATION FOR SEQ ID NO: 160:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2120 base pairs

(B) TYPE: nucleic acid

60 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 160:

5	CCCCGATAC CGCCTGACGT AGTGCCAATC ACACCTCTCG CGTCTCGGCG CCTCGGAGGC	60
	TAATGAGGAC GCCTGGCGAA ACGCAGTAAC GGATTTCGG GTGGACCTTC GCTTTACGGC	120
	TCGTGAGTTC TTCCGCCCAA CCCAGAGGAA GCGGGAGAGC AGTTTACGAC AGCGCCGGTC	180
10	GTGTTTACGG CGGCGCCCGC TGGCGCGCA TGTTCCTCT TTTCTGGTT TCTCAAGAGT	240
	GCTGCTGCTA ACGCGGTCCC CGGCACGCAC CATCTGTTGC CATCCCGGCC GGCCGAGGCA	300
15	TTGCAGATTT TGGAAGATGG CAAAGTTCAT GACACCCGTG ATCCAGGACA ACCCCTCAGG	360
	CTGGGGTCCC TGTGCGGTTT CCGAGCAGTT TCGGGATATG CCTACCAGC CGTTCAGCAA	420
	AGGAGATCGG CTAGGAAAGG TTGCAGACTG GACAGGAGCC ACATACCAAG ATAAGAGGTA	480
20	CACAAATAAG TACTCCTCTC AGTTTGGTGG TGGAAGTCAA TATGCTTATT TCCATGAGGA	540
	GGATGAAAGT AGCTTCCAGC TGGTGGATAC AGCGCGCACA CAGAAGACGG CCTACCAGCG	600
25	GAATCGAATG AGATTTGCCC AGAGGAACCT CCGCAGAGAC AAAGATCGTC GGAACATGTT	660
	GCAGTTCAAC CTGCAGATCC TGCCTAAGAG TGCCAAACAG AAAGAGAGAG AACGCATTCTG	720
	ACTGCAGAAA AAGTTCCAGA AACAATTTGG GGTAGGCAG AAATGGGATC AGAAATCACA	780
30	GAAACCCCGA GACTCTTCAG TTGAAGTTCT TAGTGATTGG GAAGTGAAAG AGGAAATGGA	840
	TTTCTCTCAG TTGATGAAGA TGCCTACTTT GGAAGTATCA GAGCCACAGG ACATTGAGTG	900
35	TTGTGGGGCC CTAGAATACT ACGACAAAGC CTTTGACCGC ATCACCACGA GGAGTGAGAA	960
	GCCACTGCGG ASATNCAAGC GCATCTTCCA CACTGTCAAC ACCACAGAGC ACCCTGTCTAT	1020
	CCGCAAGCTG GCAAAACTC AGGGGAATGT GTTTGCCACT GATGCCATCC TGGCCACGCT	1080
40	GATGAGCTGT ACCCGCTCAG TGTATTCTTG GGATATTGTC GTCCAGAGAG TTGGGTCCAA	1140
	ACTCTTCTTT GACAAGAGAG ACAACTCTGA CTTTGACCTC CTGACAGTGA GTGAGACTGC	1200
45	CAATGAGCCC CCTCAAGATG AAGGTAATTC CTTCAATTCA CCCCACAACC TGGCCATGGA	1260
	GGCAACCTAC ATCAACCACA ATTTCTCCCA GCAGTGCTTG AGAATGGGGA AGGAAAGATA	1320
	CAACTTCCCC AACCCAAACC CGTTGTGGA GGACGACATG GATAAGAATG AAATCGCCTC	1380
50	TGTTGCGTAC CGTTACCGCA GTGNAAGCT TGGAGATGAT ATTGACCTTA TTGTCCGTTG	1440
	TGAGCACGAT GCGTCATGA CTGGAGCCAA CGGGGAAGTG TCCTTCATCA ACATCAAGAC	1500
55	ACTCAATGAG TGGGATTTCA GGCAGTGTA TGGCGTTGAC TGGCGTCAGA AGCTGGACTC	1560
	TCAGCGAGGG GCTGTCAATT CCACGGAGCT GAAGAACAAC AGCTACAAGT TGGCCCGGTG	1620
60	GACCTGCTGT GCTTGTCTGG CTGGATCTGA GTACCTCAAG CTTGGTTATG TGTCTCGGTA	1680

CCACGTGAAA GACTCCTCAC GCCACGTCAT CCTAGGCACC CAGCAGTTCA AGCCTAATGA 1740
GTTTGCCAGC CAGATCAACC TGAGCGTGGA GAATGCCTGG GGCATTTTAC GCTGCGTCAT 1800
5 TGACATCTGC ATGAAGCTGG AGGAGGGCAA ATACCTCATC CTCAGGACC CCAACAAGCA 1860
GGTCATCCGT GTCTACAGCC TCCCTGATGG CACCTTCAGC TCTGATGAAG ATGAGGAGGA 1920
AGAGGAGGAG GAAGAAGAGG AAGAAGAAGA GGAAGAAACT TAAACCAGTG ATGTGGAGCT 1980
10 GGAGTTTGTC CTTCCACCGA GACTACGAGG GCCTTTGATG CTTAGTGGAA TGTGTGTCTA 2040
ACTTGCTCTC TGACATTTAG CAGATGAAAT AAAATATATA TCTGTTTAGT CTTAAAAAAA 2100
15 AAAAAAAAAA AAAAAAAAAAN 2120

20 (2) INFORMATION FOR SEQ ID NO: 161:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 900 base pairs

(B) TYPE: nucleic acid

25 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 161:

30 GGAAGCTGAA GTCCTTCCAG ACCAGGGACA ACCAGGGCAT TCTCTATGAA GCTGCACCCA 60
CCTCCACCCT CACCTGTRAC TCAGGACCAC AGAAGCAAAA GTTCTCACTC AAAGTGGATG 120
CCAAGGATGG GCGCTTGTTT AATGAGCAGA ACTTCTTCCA GCGGGCCGCC AAGCCTCTGC 180
35 AAGTCAACAA GTGAAGAAG CTGTACTCGA CCCCACTGCT GGCCATCCCT ACCTGCATGG 240
GTTTCGGTGT TCACCAGGAC AAATACAGGT TCTTGGTGTT ACCCAGCCTG GGGAGGAGCC 300
40 TTCAGTCGGC CTGGATGTC AGCCCAAAGC ATGTGCTGTG CAGAGAGGTC TGTGCTGCAG 360
GTGGCCTGCC GGCTGCTGGA TGCCCTGGAG TTCCTCCATG AGAATGAGTA TGTTCATGGA 420
AATGTGACAG CTGAAAAATAT CTTTGTGGAT CCAGAGGACC AGAGTCAGGT GACTTTGGCA 480
45 GGCTATGGCT TCGCMTCCG CTATTGCCCA AGTGGCAAAC ACGTGGCCTA CGTGAAGGC 540
AGCAGGAGCC CTCACGAGGG GGACCTTGAG TTCATTAGCA TGGACCTGCA CAAGGGATGC 600
50 GGGCCCTCCC GCCGCRGCGA CCTCCAGAGC CTGGGCTACT GCATGCTGAA GTGGCTCTAC 660
GGGTTTCTGC CATGGACAAA TTGCCTTCCC AAMAMTGAGG ACATCATGAA GCAAAAACAG 720
AAGTTTGTGTG ATAAGCCGGG GCCCTTCGTG GGACCTGCG GTCACCTGAT CAGGCCCTCA 780
55 GAGACCCTGC AGAAGTACCT GAAGGTGGTG ATGGCCCTCA CGTATGAGGA GAAGCCGCCC 840
TACGCCATGC TGAGGAACAA CCTAGAAGCT TTGCTGCAGG ATCTGCGTGT GTCTCCATAT 900
60

(2) INFORMATION FOR SEQ ID NO: 162:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1003 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
10 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 162:

GGCACGAGAT GAGGGGCACC CAGTGCTTCT AGGGCAGGCT GGGTGGTGGT CCCCTAGGTA 60
15 TCAGCCTCTC TTACTGTACT CTCGGGAAT GTTAACCTTT CTATTTTCAG CCTGTGCCAC 120
CTGTCTAGGC AAGCTGGCTT CCCCATTTGGC CCCTGTGGGT CCACAGCAGC GTGGCTGCCC 180
CCCAGGGCCA CCGCTTCTTT CTTGATCCTC TTTCCTTAAC AGTGACTTGG GCTTGAGTCT 240
20 GGCAAGGAAC CTGCTTTTA GCTTACCAC CAAGGAGAGA GGTGACATG ACCTCCCCGC 300
CCCTCACCA AGGCTGGGAA CAGAGGGGAT GTGGTGAGAG CCAGGTTCTT CTGGCCCTCT 360
25 CCAGGGTGT TTCCACTAGT CACTACTGTC TTCTCCTTGT AGCTAATCAA TCAATATTCT 420
TCCCTTGCTT GTGGGCAGTG GAGAGGCTGC TGGGTGTACG CTGCACCTGC CCACTGAGTT 480
GGGAAAGAG GATAATCAGT GAGCACTGTT CTGCTCAGAG CTCCTGATCT ACCCCACCCC 540
30 CTAGGATCCA GGACTGGGTC AAAGCTGCAT GAAACCAGGC CCTGGCAGCA AACCTGGGAA 600
TGGCTGGAGG TGGGAGAGAA CCTGAATTC TCTTTCCCTC TCCCTCTCC AACATTACTG 660
35 GAACTCTATC CTGTTAGGAT CTTCTGAGCT TGTTCCTG CTGGGTGGGA CAGAGGACAA 720
AGGAGAAGGG AGGGTCTAGA AGAGGCAGCC CTTCTTTGTC CTCTGGGGTA AATGAGCTTG 780
ACCTAGAGTA AATGGAGAGA CAAAAGCCT CTGATTTTAA ATTTCCATAA AATGTTAGAA 840
40 GTATATATAT ACATATATAT ATTTCTTTAA ATTTTIGAGT CTTTGATATG TCTAAAAATC 900
CATTCCTCT GCCCTGAAGC CTGAGTGAGA CACATGAAGA AAAGTGTGTT TCATTTAAAG 960
45 ATGTTAATTA AATGATTGAA ACTTGAAAAA AAAAAAAAAA AAA 1003

50 (2) INFORMATION FOR SEQ ID NO: 163:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2196 base pairs
(B) TYPE: nucleic acid
55 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 163:

60 AAGAAGCGGC ACACGGATGT GCAGTTCTAC ACAGAAGTGG GAGAGATAAC CACGGACTTG 60

	GGGAAACATC AGCATATGCA TGACCGAGAT GACCTCTATG CTGAGCAGAT GGAACGAGAA	120
5	ATGAGGCACA AACTGAAAAC AGCCTTTAAA AATTTTCATTG AGAAAGTAGA GGCTCTAACT	180
	AAGGAGGAAC TGGAAATTGA AGTGCCTTTT AGGGACTTGG GATTTAACGG AGCTCCCTAT	240
	AGGAGTACCT GCCTCCTTCA GCCCACTAGT AGTGCCTTGG TAAATGCTAC GGAATGGCCA	300
10	CCTTTTGTGG TGACATTGGA TGAGGTAGAG CTGATCCACT TTRAGCGGGT CCAGTTTCAC	360
	CTGAAGAACT TTGATATGGT AATCGTCTAC AAGGACTACA GCAAGAAAGT GACCATGATC	420
15	AACGCCATTCT CTGTAGCCTC TCTTGACCCC ATCAAGGAAT GGTGAATTC CTGCGACCTG	480
	AAATACACAG AAGGAGTACA GTCCTCAAC TGGACTAAAA TCATGAAGAC CATTTGTGAT	540
	GACCCTGAGG GCTTCTTCGA ACAAGGTGGC TGGTCTTTCC TGGAGCCTGA GGGTGAGGGG	600
20	AGTGATGCTG AAGAAGGGGA TTCAGAGTCT GAAATTGAAG ATGAGACTTT TAATCCTTCA	660
	GAAGATGACT ATGAAGAGGA AGAGGAGGAC AGTGATGAAG ATTATTTCATC AGAAGCAGAA	720
25	GAGTCAGACT ATTCTAAGGA GTCATTGGGT AGTGAAGAAG AGAGTGGAAG GGATTGGGAT	780
	GAACTGGAGG AAGAAGCCCG AAAAGCGGAC CGAGAAAGTC GTTACGAGGA AGAAGAAGAA	840
	CAAAGTCGAA GTATGAGCCG GAAGAGGAAG GCATCTGTGC ACAGTTCCGG CCGTGGCTCT	900
30	AACCGTGGTT CCAGACACAG CTCTGCACCC CCCAAGAAA AGAGGAAGTA ACTTCTGAAC	960
	TTTGGCCCTG AGCTCCATTCTTCTCCAGC CAACCCCTGA AAATTTTACA TGACATAGAA	1020
35	ACTGTATTTT TCCTTTTGGT TTCATTGAA GTTTTGCCAT TTGTGTTTAT GGGTTTAGGG	1080
	GGCCATTTGT GTGGACCAAT CTACTCGGGG AATTCCAGGC CCACCAGGAC ACGTGCCAAT	1140
	GGCCCCATTC AGATGGCAAG GGAGGAGGTG TTCTTGAAGA CAGGAGGAGG CTCCCGCTGT	1200
40	TAATAAATAT TGTTCATTCTTCTCTCTC CTGTACCTT CTGCCAAGAC ATTGATGGCT	1260
	TCTGACATCT TATTTGGTGT CTCAAAGCTG TATTTCCAAG ACAGTGGTAC AAGGTGACCC	1320
45	TTAATTACCC GTATCATGGT TCTTGACCAG CACATTCAAT CCTCCAACCT ACCCTACTGC	1380
	CATGACCTTC CGCACATCTC TAAGTTTAT CTTTGCAATA CTCAAGGTTT TCGGAAATTT	1440
	GCTAATGGTT GTGATAAACC ATACAGCTTG AGCCAGTGAG GCAGATTGGG CTGGTGCCTT	1500
50	CGTCTGAGTT TTCTGCTTT CTGCCTCGT GCAGATTCTG AGGTATATCT GCTGCCTTGG	1560
	AAGACATAAG AAGCAGTGAT ACTCCCTGGC TCGGTATTTT TCTCCATACA ATGCACACAT	1620
55	GGTACAATGA TAGAAGGCAA AATTGCCACT GTCTTCTTTT TTTTCTCATA TATCTAAGGA	1680
	AGATATATCA GGTGTGCCT CATGTACCGC TTCTAGTGAA ATGTAGAGGA AGGCTCAAAG	1740
	GAGTCAACAT TTAGATCTGG AAGGACAAG TCATGCCTTG GGCCTAGAAT ACCCTGATGA	1800
60	GAAAAGAGAA GAGGAAGGGA GGCCATATCT ACAACANCAN CCTCTCGGCA CTGCTGCTCC	1860

TTATTTTAAC TTGTCTTGC ATTGTCCTGT ATTTATCACA GTTCTGTG AACAGCTTTT 1920
CAAGTATTTG GGGAGTTTAT CTGCCATCC TCCCCTTCTG GTTCTCTGCA CCCACCTGTC 1980
5 CCACTGCAGT TCCTTCCGTG CTCTGTGACT TTAAGAGAAG AAGGGGGGAG GGGTCCCGGA 2040
TTTATGTTT GTTGTMTT TCTCCTTAGC AGTAGGACTT GATATTTTCA ATTTTGAAG 2100
10 AACTAAAAGA TGAATAAACT GGGTTTTTTT TGTGTGTTGT TTTTGTAAAA AAAAAAAAAA 2160
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAA 2196

15

(2) INFORMATION FOR SEQ ID NO: 164:

(i) SEQUENCE CHARACTERISTICS:
20 (A) LENGTH: 1945 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 164:

GCACAGAGTC GGGCGGACCG ACAGGGAGAG GAGGAGAGGG GGTCTGCCG CGGCCGCTAC 60
CCAGAAAGCCA GCGGACGGCA GCACGGAGTG GGCTGTCCCC GAGCCCAGCC CCGAGCGAGC 120
30 CCCCCCCCCG CCCCCGMAGG ACGCGCCTYC CAGCCAGCCC GACTVCTAGG AGGAGGGGAG 180
GCGGGAAGC AGCTCAAGCC TCACCCACCG CCCTGCCCCC AGCCCCGCCA CTCCCAGGCT 240
35 CCTCGGGACT CGGCGGGTCC TCCTGGGAGT CTCGGAGGGG ACCGGCTGTG CAGACGCCAT 300
GGAGTTGGTG CTGGTCTTCC TCTGCAGCCT GCTGGCCCC ATGGTCTCTG CCAGTGCAGC 360
TGAAAAGGAG AAGGAAATGG ACCCTTTTCA TTATGATTAC CAGACCCTGA GGATTGGGG 420
40 ACTGGTGTTC GCTGTGGTCC TCTTCTCGGT TGGGATCCTC CTTATCTTAA GTCGCAGGTG 480
CAAGTGCAGT TTCAATCAGA AGCCCCGGC CCCAGGAGAT GAGGAAGCCC AGGTGGAGAA 540
45 CCTCATCACC GCCAATGCAA CAGAGCCCCA GAAAGCAGAG AACTGAAGTG CAGCCATCAG 600
GTGGAAGCCT CTGGAACCTG AGGCGGCTGC TTGAACCTTT GGATGCAAAT GTCGATGCTT 660
AAGAAAACCG GCCACTTCAG CAACAGCCCT TTCCCCAGGA GAAGCCAAGA ACTTGTGTGT 720
50 CCCCCACCCT ATCCCCTCTA ACACCATTC TCCACCTGAT GATGCAACTA ACACTTGCCT 780
CCCCACTGCA GCCTGCGGTC CTGCCACCT CCCGTGATGT GTGTGTGTGT GTGTGTGTGT 840
55 GTGACTGTGT GTGTTTGCTA ACTGTGGTCT TTGTGGCTAC TTGTTTGTGG ATGGTATTGT 900
GTTTGTAGT GAACTGTGGA CTCGCTTTC CAGGCAGGGG CTGAGCCACA TGGCCATCTG 960
CTCCTCCCTG CCCCCGTGGC CCTCCATCAC CTTCTGCTCC TAGGAGGCTG CTTGTGCCCC 1020
60

GAGACCAGCC CCCTCCCCTG ATTTAGGGAT GCGTAGGGTA AGAGCACGGG CAGTGGTCTT 1080
CAGTCGTCTT GGGACCTGGG AAGGTTTGCA GCACCTTGTC ATCATTCTTC ATGGACTCCT 1140
5 TTCACTCCTT TAACAAAAAC CTTGCTTCCT TATCCACCT GATCCCAGTC TGAAGGTCTC 1200
TTAGCAACTG GAGATACAAA GCAAGGAGCT GGTGAGCCCA GCGTTGACGT CAGGCAGGCT 1260
10 ATGCCCTTCC GTGGTTAATT TCTTCCCAGG GGCTTCCACG AGGAGTCCCC ATCTGCCCCG 1320
CCOCTTCACA GAGCGCCCGG GGATTCCAGG CCCAGGGCTT CTA CTCTGCC CCTGGGGAAT 1380
GTGTCCCCTG CATATCTTCT CAGCAATAAC TCCATGGGCT CTGGGACCCT ACCCTTCCA 1440
15 ACCTTCCCTG CTTCTGAGAC TTCAATCTAC AGCCAGCTC ATCCAGATGC AGACTACAGT 1500
CCCTGCAATT GGGTCTCTGG CAGGCAATAG TTGAAGGACT CCTGTTCCGT TGGGGCCAGC 1560
ACACCGGAT GGATGGAGGG AGAGCAGAGG CCTTTGCTTC TCTGCCTACG TCCCCTTAGA 1620
20 TGGGCAGCAG AGGCAACTCC CGCATCCTTT GCTCTGCCTG TCRGTGGTCA GAGCGGTGAG 1680
CGAGGTGGGT TGGAGACTCA GCAGGCTCCG TGCAGCCCTT GGGAACAGTG AGAGGTGAA 1740
25 GGTCATAACG AGAGTGGGAA CTCAACCCAG ATCCCGCCCC TCCTGTCTC TGTGTTCCCG 1800
CGGAAACCAA CCAACCGTG CGCTGTGACC CATTGCTGTT CTCTGTATCG TGATCTATCC 1860
30 TCAACAACAA CAGAAAAAAG GAATAAATA TCCTTTGTTT CTTAGTGAAA AAAAAAAAAA 1920
AAAAAAAAA AAAAAAAAAA CTCGA 1945

35

(2) INFORMATION FOR SEQ ID NO: 165:

(i) SEQUENCE CHARACTERISTICS:

40

- (A) LENGTH: 2933 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 165:

45

GGGTCGACCC ACGGTCCGG CAGCGTCTG TTGAGTCGTT GCTGCGCTG CCCCCCTCCG 60
GATCAGGAGC CAGTGTATAC CGCCGCCCA CCGCCTGGT GCCGCTAGAG GAAACGAGAA 120
50 GGAGGCCGCC TCGGTTTGT CGCCGAGCT CGCCMCYGY CYGGRAGAGC CGAGCCCCG 180
CCCAGTCGGT CGCTGCCAC CSCTCGTAGC CGTTACCCG GGGCCGCCAC AGCCGCCGGC 240
CGGGAGAGGC GCGGCCATG GCTCTGGAG CCGATTCAA AGGTGATGAC CTATCAACAG 300
55 CCATTCTCAA ACAGAAGAAC CGTCCCAATC GGTAAATTGT TGATGAAGCC ATCAATGAGG 360
ACAACAGTGT GGTGTCCTG TCCAGCCCA AGATGGATGA ATTGCAGTTG TTCCGAGGTG 420
60 ACACAGTGTT GCTGAAAGGA AAGAAGAGAC GAGAAGCTGT TTGCATCGTC CTTCTGATG 480

	ATACTGTGTC TGATGAGAAG ATTCGGATGA ATAGAGTTGT TCGGAATAAC CTTCTGTGTAC	540
	GCCTAGGGGA TGTCATCAGC ATCCAGCCAT GCCCTGATGT GAAGTACGGC AAACGTATCC	600
5	ATGTGCTGCC CATTGATGAC ACAGTGAAG GCATTACTGG TAATCTCTTC GAGGTATACC	660
	TTAAGCCGTA CTTCTGGAA GCGTATCGAC CCATCCGAA AGGAGACATT TTTCTTGTC	720
10	GTGGTGGGAT GCGTGCTGTG GAGTTCAAAG TGGTGGAAAC AGATCCTAGC CCTTATTGCA	780
	TTGTGTCTCC AGACACAGTG ATCCACTGCG AAGGGGAGCC TATCAAACGA GAGGATGAGG	840
	AAGAGTCCIT GAATGAAGTA GGGTATGATG ACATTGGTGG CTGCAGGAAG CAGCTAGCTC	900
15	AGATAAAGGA GATGGTGGAA CTGCCCTGA GACATCCTGC CCTCTTTAAG GCAATTGGTG	960
	TGAAGCCTCC TAGAGGAATC CTGCTTTACG GACCTCCTGG AACAGGAAAG ACCCTGATTG	1020
20	CTCGAGCTGT AGCAAATGAG ACTGGAGCCT TCTTCTTCTT GATCAATGGT CCTGAGATCA	1080
	TGAGCAAATT GGCTGGTGAG TCTGAGAGCA ACCTTCGTAA AGCCTTTGAG GAGGCTGAGA	1140
	AGAATGCTCC TGCCATCATC TTCATTGATG AGCTAGATGC CATCGCTCCC AAAAGAGAGA	1200
25	AAACTCATGG CGAGGTGGAG CGGCGCATTG TATCACAGTT GTTGACCCTC ATGGATGGCC	1260
	TAAAGCAGAG GGCACATGTG ATGTGTTATGG CAGCAACCAA CAGACCCAAC AGCATTGACC	1320
30	CAGCTCTACG GCGATTGGT CGCTTTGACA GGGAGGTAGA TATTGGAATT CCTGATGCTA	1380
	CAGGACGCTT AGAGATTCTT CAGATCCATA CCAAGAACAT GAAGCTGGCA GATGATGTGG	1440
	ACCTGGAACA GTAGCCAATG AGACTCACGG GCATGTGGGT GCTGACTTAG CAGCCCTGTG	1500
35	CTCAGAGGCT GCTCTGCAAG CCATCCGCAA GAAGATGGAT CTCATTGACC TAGAGGATGA	1560
	GACCATTGAT GCCAGGTCA TGAATCTCT AGCAGTTACT ATGGATGACT TCCGGTGGGC	1620
40	CTTGAGCCAG AGTAACCCAT CAGCACTGCG GGAAACCGTG GTAGAGGTGC CACAGGTAAC	1680
	CTGGAAGAC ATCGGGGGCC TAGAGGATGT CAAACGTGAG CTACAGGAGC TGGTCCAGTA	1740
	TCCTGTGGAG CACCCAGACA AATTCTGAA GTTTGGCATG ACACCTTCCA AGGGAGTTCT	1800
45	GTCTATGGA CCTCTGGCT GTGGGAAAAC TTTGTTGGCC AAAGCCATTG CTAATGAATG	1860
	CCAGGCAAAC TTCATCTCCA TCAAGGGTCC TGAGCTGCTC ACCATGTGGT TTGGGGAGTC	1920
50	TGAGGCCAAT GTCAGAGAAA TCTTTGACAA GGCCCGCAA GCTGCCCCCT GTGTGCTATT	1980
	CTTTGATGAG CTGGATTGCA TTGCCAAGGC TCGTGGAGGT AACATTGGAG ATGGTGGTGG	2040
	GGCTGCTGAC CGAGTCATCA ACCAGATCCT GACAGAAATG GATGGCATGT CCACAAAAA	2100
55	AAATGTGTTT ATCATTTGGC CTACCAACCG GCCTGACATC ATTGATCCTG CCATCCTCAG	2160
	ACCTGGCCGT CTTGATCAGC TCATCTACAT CCCACTTCCT GATGAGAAGT CCCGTGTTGC	2220
60	CATCCTCAAG GCTAACCTGC GCAAGTCCCC AGTTGCCAAG GATGTGGACT TGGAGTTCCT	2280

GGCTAAATG ACTAATGGCT TCTCTGGAGC TGACCTGACA GAGATTGCC AGCGTCTTG 2340
CAAGCTGGCC ATCCGTGAAT CCATCGAGAG TGAGATTAGG CGAGAACGAG AGAGGCAGAC 2400
5 AAACCCATCA GCCATGGAGG TAGAAGAGGA TGATCCAGTG CCTGAGATCC GTCGAGATCA 2460
CTTTGAAGAA GCCATGCGCT TTGCGCGCCG TTCTGTCACT GACAATGACA TTCGGAAGTA 2520
10 TGAGATGTTT GCCCAGACCC TTCAGCAGAG TCGGGGCTTT GGCAGCTTCA GATTCCCTTC 2580
AGGGAACCAG GGTGGAGCTG GCCCCAGTCA GGCAGTGA GCGGCACAG GTGGCAGTGT 2640
ATACACAGAA GACAATGATG ATGACCTGTA TGGCTAAGTG GTGGTGGCCA GCGTGCAGTG 2700
15 AGCTGGCCTG CCTGGACCTT GTTCCCTGGG GTGGGGGCG CTGCCCAGG AGAGGGACCA 2760
GGGTGCGCC CACAGCCTGC TCCATTCTCC AGTCTGAACA GTTCAGCTAC AGTCTGACTC 2820
20 TGGACAGGGG GTTCTGTG CAAAAATACA AAACAAAGC GATAAAATAA AAGCGATTTT 2880
CATTTGGTAA AAAAAAAAAA AAAAAAAAAAT CCGGGGGGGG GCCCGAACCA TTT 2933

25

(2) INFORMATION FOR SEQ ID NO: 166:

(i) SEQUENCE CHARACTERISTICS:
30 (A) LENGTH: 2243 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 166:

TCGGAGAGCC GCGGGGCGNG CGCCTCTCGG CCAGGAAGCG CCTCTTGGAC GCGTGTNACC 60
GATGCCCAGA AGTGGCCTTG GGTGGGGAT CACCATAGCT TTTCTAGCTA CGCTGATCAC 120
40 GCAGTTTCTC GTGTATAATG GTGTCTATCA GTATACATCC CCAGATTTC TCTATATTCC 180
TTCTTGGCTC CCTTGTATAT TTTTCTCAGG AGCGCTCAG GTGGGGAACA TAGGACGACA 240
45 GTTAGCTATG GGTGTTCCCTG AAAAGCCCCA TAGTGATTGA GTCTTCAAAA CCACCGATTTC 300
TGAGAGCAAG GAAGATTTTG GAAGAAAATC TGA CTGTGGA TTATGACAAA GATTATCTTT 360
TTTCTTAAGT AATCTATTTA GATCGGGCTG ACTGTACAAA TGACTCCTGG AAAAACTCT 420
50 TCACCTAGTC TAGAATAGGG AGGTGGAGAA TGATGACTTA CCCTGAAGTC TTCCCTTGAC 480
TGCCCGCACT GCGCCTGTC TGTGCCCTGG AGCATCTGC CCAGGCTACG TGGGTTCAGG 540
55 CAGGTGGCAG CTTCCTCAAGT ATTCGATTTC ATTCATGTGA TTAAACAAG TTGCCATATT 600
TCAAAGCCTT GAACTAAGAC TCAATTACCA ACCCGCAGTT TTGTGTCACT GCCCAAAGGA 660
GGTAGGTGTA TGGTGCTTAA CAAACATGAA GTATGGTGTA ATAGGAATAA TATTTATCCA 720
60

	AAAGATTTTT AAAAATAGGG CTGTGTTTAA AAAAAAAAC AAAACARGAA AAGCAGCAGT	780
	GATTATAGAG AGGTCACACT CTAAGTGGGG TCGCGGCGTG GCCACGCTTC ACGGTCACGC	840
5	TCGTCCGTCC TGCAGTGGCG TGTITACATG GTCACACGTG TGTGTATCAC CAGTGGGTCA	900
	ACTGCTTGTC ATTCCTCCCG TGGCAGTTTG TGTAGACAAT CTACTGAGC AAAAGGCAAT	960
10	GAAAAGTCTT GGTTCACACA CTGCGATATA TTGGAATTTT CACCTCAGTT TATGAAGTTT	1020
	ATTTGCAAAT CCATAGTCAT CTAAGAATGA ATACCTGTCT GCCATGTATT TCAATCTTAG	1080
	TGAGCCAAAA TTGTTGTGTT GTTACTACAG AATAGAGATG ACTGTTTTTT GCCACAGCCC	1140
15	TATGGRATTT GCAATCTGTG ATTGCCTTGT AAAAAGGAGA GTGCATATGG CACTGCATTA	1200
	AACGTGTGGT GTTCTAGTC AATGATATTG GTGAGCACAA TGTATTCATT TAATGGCATA	1260
20	GACCATACCA GACCTAATTT GCAAGTATTG GGTCTTAAAC TTCAAGTGCA ATGTATATGA	1320
	AAACCAATCT GAGCCTGTGA TCTCTTAAAT ATTTATTTTT TTTAACGTGT GAGATGTTTG	1380
	AGAGAAGGTT CTCCATTCAT TTCAGTGCTG CCTGGAGGAA ACTCGGCAAT GATTTCTTTC	1440
25	AGTTGTGAAG TTCCTTTCGT GTTACACCCT CCACTGAACC CTCAACCTTC GAAATACTCC	1500
	AGTTTGTGG GTTTGGTCAT TTTTACTTAT AAATTTACCT TTTTGTATTT TGCAATTTAC	1560
30	ATGTGTTTGG TTTGTTTAA ATCTGTGAA AGTGGCTTGA TTAAAAGACT CCTTTTAAAT	1620
	GGAAGCCACC AGTCAGCAGA ATGGAAGCTT AGAGGAACCT GCCTGTGAGC GCTGGTCTTT	1680
	GTGTTTGGTT TTGTGATGTA ACGATCTTTG CTGGGGTTTT TTGCTTTGTT TTGAGGGAAA	1740
35	TGCTTTGGAG TAAATTTTAA GTTCTTGAG TTAATTTGTT TTACAGGAAT TTTGTTTTTT	1800
	AAAAAATAG GATCATCTG AACTTTGGAA TGACCCCTT ATATATTTTC TGAAAATGAA	1860
40	AACAGTTACA TGAAAAAAT TTCCAATGAA GATGTCAGCA TTTTATGAAA AACCAGAAGT	1920
	TATTAGATGA AAGCAGCGAG TGAATCTTTA AACAGACTT GATCACGCAC ACACAATAAG	1980
	TCTTTCTCTC CGAAACCGGA AGTAAATCTA TATCTGTTAG AAATAATGTA GCCAAAAGAA	2040
45	TGTAAATTTG AGGATTTTTT TGCCAATAGT TTATAGAAAA TATATGAACC AAAGTGATTT	2100
	GAGTTTGTA AAATGTAAAA TAGTATGAAC AAAATTTGCA CTCTACCAGA TTTGAACATC	2160
50	TAGTGAGGTT CACATTCATA CTAAGTTTTC AACATTGTGT TCTTTTGTGA TTCATTTTTT	2220
	ACTTTTATTA AAGGTTCAAA ACC	2243

55

(2) INFORMATION FOR SEQ ID NO: 167:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1816 base pairs

(B) TYPE: nucleic acid

60

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 167:

5	GGTGGGNAGC TTTNAATTTC CCCTTACWGG GCGCTNTAA GGGGAAACCT TCCCAGGAATT	60
	TTCCGGGTCGA CCCACGCGTC CGGCCAGCCT AGGAGAAGAA GTTCGTAGTC CCAGAGGTGA	120
10	GGCAGGAGGC GGCAGTTTCT GCGGGGTGAG GCGGAGCTG AAGTGACAGC GGAGGCGGAA	180
	GCAACGGTCG GTGGGGCGGA GAAGGGGGCT GGGCCAGGA GGAGGAGGAA ACCCTTCCGA	240
15	GAAAACAGCA ACAAGCTGAG CTGCTGTGAC AGAGGGGAAC AAGATGGCGG CGCCGAAGGG	300
	GAGCCTCTGG GTGAGGACCC AACTGGGGCT CCCGCCGCTG CTGCTGCTGA CCATGGCCTT	360
	GGCCGGAGGT TCGGGGACCG CTTCGGCTGA AGCATTTGAC TCGGTCITGG GTGATACGGC	420
20	GTCTTGCCAC CGGGCTGTC AGTTGACCTA CCCCTTGAC ACCTACCTA AGGAAGAAGA	480
	GTGTACGCA TGTACAGAG GTTCAGGCT GTTTTCAATT TGTAGTTTG TGGATGATGG	540
	AATTGACTTA AATCGAATA AATTGGAATG TGAATCTGCA TGTACAGAAG CATATTCCCA	600
25	ATCTGATGAG CAATATGCTT GCCATCTTGG KTGCCAGAAT CAGCTGCCAT TCGCTGAAT	660
	GAGACAAGAA CAACTTATGT CCTGATGCC AAAAATGCAC CTACTCTTTC CTCTAACTCT	720
30	GGTGAGGTCA TTCTGGAGTG ACATGATGGA CTCGCACAG AGCTTCATAA CCTCTTCATG	780
	GACTTTTTTAT CTTCAAGCCG ATGACGGAAA AATAGTTATA TTCCRGCTA AGCCAGRAA	840
35	TCCCAGGTAC GCACCACATT TGGAGCCAGG AGCCCTACCA AATTGTGRGR RAWCMCTCT	900
	AAGCAAAATG TCCNTCAKMT CGSMAATGAG AAATTCACAA GCGCACAGGA ATTTTCTTGA	960
	AGATGGAGAA AGTGATGGCT TTTTAAGATG CCTCTCTCTT AACTCTGGGT GGATTTTAAC	1020
40	TACAACTCTT GTCCTCTCGG TGATGGTATT GCTTTGGATT TGTTGTGCAA CTTGTGCTA	1080
	CACGCTGTTG GACGCAGTAT AGTTTCCCTC TGAGAAGCTG AGTATCTATG GTGACTTGGA	1140
	GTTTATGAAT GAACAAAAGC TAAACAGATA TCCAGCTTCT TCTCTTGTGG TTGTTAGATC	1200
45	TAAAAC TGAA GATCATGAAG AAGCAGGGCC TCTACCTACA AAAGTGAATC TTGCTCATTC	1260
	TGAAATTTAA GCATTTTCT TTTAAAAGAC AAGTGAATA GACATCTAAA ATTCCACTCC	1320
50	TCATAGAGCT TTTAAAATGG TTTCATTGGA TATAGGCCTT AAGAAATCAC TATAAAATGC	1380
	AAATAAAGTT ACTCAAATCT GTGAAAAAAA AAAAAAAAAA AAAAAAAAC TCGAGGGGGG	1440
	GCCCGTTACC AAKTCGCCCT ATWGTGADTB GTATTMTTAT TTTACTAATA TCTGTAGCTA	1500
55	TTTTGTTTTT KGCTTKGGTT ATKGTTTTTY TCCCTTYTCT WAGCTATRAG CTGATCATKG	1560
	CYSCTTCTCA CCTCTGCCA TGATACTGTC AGTTACCTTA GTTAACAAGC TGAATATTTA	1620
60	GTAGAAATGA TGCTTCTGCT CAGGAATGGC CCACAAATCT GTAATTTGAA ATTTAGCAGG	1680

AAATGACCTT TAATGACACT ACATTTTCAG GAACTGAAAT CATTAAAATT TTATTGAAT 1740
AATTATGTGC TGAAAAAAAA AAAAAAAAAA AMWMRARASK RRWWACTCGA GGGGGGGCCC 1800
5 GGTACCCNAT TCGCCG 1816

10

(2) INFORMATION FOR SEQ ID NO: 168:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 945 base pairs
15 (B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 168:

AGAAACCGTT GATGGGACTG AGAAACCAGA GTTAAACCT CTTTGAGACT TCTGAGGACT 60
CAGCTGGAAC CAACGGGCAC AGTGGCAAC ACCATCAACT TCTCCAAGC AGAGAAACCC 120
25 GAACCCACCA ACCAGGGCA GGATAGCCTG AAGAAACATC TACACGCAGA AATCAAAGTT 180
ATTGGGACTA TCCAGATCTT GTGTGGCATG ATGGTATTGA GCTTGGGGAT CATTTTGGCA 240
TCTGCTTCCT TCTCTCCAAA TTTTACCCAA GTGACTTCTA CACTGTTGAA CTCTGCTTAC 300
30 CCATTCATAG GACCCCTTTT TTTTATCATC TCTGGCTCTC TATCAATCGC CACAGAGAAA 360
AGGTTRACCA AGCTTTTGGT GCATAGCAGC CTGGTTGGAA GCATTCTGAG TGCTCTGTCT 420
35 GCCCTGGTGG GTTTCATTAT CCTGTCTGTC AAACAGGCCA CCTTAAATCC TGCCTCACTG 480
CAGTGTGAGT TGGACAAAAA TAATATACCA ACAAGAAGTT ATGTTTCTTA CTTTATCAT 540
GATTCACTTT ATACCACGGA CTGCTATACA GCCAAAGCCA GTCGGCTGG AWCTCTCTCT 600
40 CTGATGCTGA TTTGCACTCT GCTGGAATTC TGCCTAGCTG TGCTCACTGC TGTGCTGCGG 660
TGGAACAGG CTTACTCTGA CTCCCTGGG AGTGTACTTT TCCTGCCTCA CAGTTACATT 720
45 GGTAATTCTG GCATGTCCTC AAAAATGACT CATGACTGTG GATATGAAGA ACTATTGACT 780
TCTTAAGAAA AAAGGGAGAA ATATTAATCA GAAAGTTGAT TCTTATGATA ATATGGAAAA 840
50 GTTAACCATT ATAGAAAAGC AAAGCTTGAG TTTCTTAAAT GTAAGCTTTT AAAGTAATGA 900
ACATTAAAAA AAACCATTTAT TTCACTGTCA TTTAAAGATA ATGTG 945

55

(2) INFORMATION FOR SEQ ID NO: 169:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 902 base pairs
60 (B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 169:

5
GGCAGAGCCA CAGGAAGGAT GAGGAAGACC AGGCTCTGGG GGCTGCTGTG GATGCTCTTT 60
GTCTCAGAAC TCCGAGCTGC AACTAAATTA ACTGAGGAAA AGTATGAACT GAAAGAGGGG 120
10 CAGACCCTGG ATGTGAAATG TGA CTACTACG CTAGAGAAGT TTGCCAGCAG CCAGAAAGCT 180
TGGCAGATAA TAAGGGACGG AGAGATGCCC AAGACCCTGG CATGCACAGA GAGGCCTTCA 240
AAGAATTCCT ATCCAGTCCA AGTGGGGAGG ATCATACTAG AAGACTACCA TGATCATGGT 300
15 TTACTGCGCG TCCGAATGGT CAACCTTCAA GTGGAAGATT CTGGACTGTA TCAGTGTGTG 360
ATCTACCAGC CTCCTAAGGA GCCTCACATG CTGTTCTGATC GCATCCGCTT GGTGGTGACC 420
20 AAGGGTTTTT CAGGGACCCC TGGCTCCAAT GAGAATTCTA CCCAGAATGT GTATAAGATT 480
CCTCCTACCA CCACTAAGGC CTGTGTGCCC CTCTATACCA GCCCCAGAAC TGTGACCCAA 540
GCTCCACCCA AGTCAACTGC CGATGTCTCC ACTCCTGACT CTGAAATCAA CCTTACAAAT 600
25 GTGACAGATA TCATCAGGT TCCGGTGTTC AACATTGTCA TTCTCCTGGC TGGTGGATTG 660
CTGAGTAAGA GCCTGGTCTT CTCTGTCTG TTTGCTGTCA CGCTGAGGTC ATTGTGACCC 720
30 TAGGCCACG AACCCACGAG AATGTCCTCT GACTTCCAGC CACATCCATC TGGCAGTTGT 780
GCCAAGGGAG GAGGGAGGAG GTAAAAGCA GGGAGTTAAT AACATGAATT AAATCTGTAA 840
TCACCRGCTA AAAAAAAAAA AAAAAAACN CGANCCTNGG TTTTCAGCTC CATCAGCTCC 900
35 TT 902

40
(2) INFORMATION FOR SEQ ID NO: 170:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1883 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 170:

50
AGAAAACAAC TGAAAACCA CATTTTTCTA CATACAGCTG GGGAGGTAGC TGAGAACTTG 60
GCACTGCGCA CACATACTAG GTTGAAAGAG AGTTGAGGAA ACCAGAAGGC CAAGTGGATC 120
55 TGCTGGCAAA CCCTGAACCT GTCTCCTGCG CTGTCTCTAC AGTTCTGAAG TTGAAAATCC 180
TTTTCATGCC TAGCATCTGC TTGAGTTATA AACCCCAAGG CAGCCATGTC ATAGACTAGT 240
GTTTACTCTT GTTTTGACTT TGTTTTAATG CTTCCTAAGA CCCAAGTGCC TCCTGCTGTT 300
60

	TCCTCCTTTG TGGTAGCCTC TGGCCATCTG GGACCTCAAT CCCAGCTTT CCCACTTTCA	360
	GCAGTCCTTT GCTCTCTTTG CTTCTACCTC AAATAGCCCC AGGAGTGGGC TTTAGTCTCC	420
5	AATATGGAGC ATYTCAAGCT TCTCCTGGGG GATGGGGATT GGGATGGGCA GAATCTGTTT	480
	TGGWCTCCG GGTATTTTCC AGTGGGTGTA AAAGCAGAGC TGGGCCTTTC CCTCTCTTAT	540
10	CCCTGAGGGT GGTAAGAAG GACTGTATCT ACACCTGTTT TTTCCCTACCT TCTCTTTTGT	600
	TAGGGAGGCC TCATTTCTAAG TTCTCTAAGA GAGTCCTTGG CTAAAGCTG TAGCAAGGGT	660
	GTGCTAGGTG GGGGATTTGG AGCAAAACCG TCGAGTAGGC ATGATACTGG TATGGAGTGG	720
15	GCCTGCAAAA TCAGACAGAA ATGGCTTGAG AAGCCGACAG GGAGCATGCC TGCTCTCAG	780
	TGATAGAGTA TGGGAGGGAC CTCCTAGCT TGGAAAATGA GAATTGAAGG GGTATGAAC	840
20	AAATAGGATG CCTAGTTGAG GATGTTCCCA AAGTTTGTTC CAATCTTATC ATTAGTAGAT	900
	TTTATAAGCC ACAGAGACAA ACCAGAAACG GAATAATGTT ACTTTGGATG CTTTATTTTT	960
	TTGTTCTAGG TGTGGCTTTG TACATGCAGA AGAATGCTAT ATGCTGCACA TTTTGCCTTT	1020
25	AAAGTCTTAC GACTTTCCCC ATTTTAGTCT AATGGGAAGA TACAGATGTG CAAGTCTGCT	1080
	TTTTTGTTTT TTGTTATTAT TTTTTTTTTT TTGCTCTGTG TTATGGACAT TTTCAGACAT	1140
30	GCACAGAAGT GGAGAGGATG GTCCTTGGAC CCCATGTGTC CATCACCTAG CTGCATCACT	1200
	TATCAGCTAT GGTCAACCTG GTTTCATCTG TATCTCTCTC TTTTCACCTG TATTGTTTAT	1260
	TGAAAATCCA AGACACTATG CCAATGCAAC CGTGACTACT TTGGGAGATT GGTAGTCTCT	1320
35	TTTGATGGTG ATAGTGATGG GGTGCACTAT CATAATCACA TCAGGTCTGC TTTTGTCTTT	1380
	TAATGTTAAC TAATGAAGTT CCAGAGATGG GCCTTAGAAA TGTGTTTAA GAATTAACAA	1440
40	GGAGTCTCAA AAAGAAATGA GAGGGATGCT TCCTTTCCCC TTGCATCTAC AAAACAAGAG	1500
	AGAGACTGTT CTGTTGTAAA ACTCTTTCAA AAATCTGAT ATGGTAAGGT ACTTGAGACC	1560
	CTTCACCAGA ATGTCAATCT TTTTTTCTGT GTAACATGGA AACTTGTGTG ACCATTAGCA	1620
45	TTGTTATCAG CTTGTACTGG TCTCATAACT CTGGTTTGG AAGAATAATT TGGAAATTGT	1680
	TGCTGTGTTT TGTGAAAATA ACCTCCCCAA AATAATTAGT AACTGGTTGT TCTACTTGGT	1740
50	AATTTGACAC CCTGTTAATA ACGCAATTAT TTCTGTGTTT TTAAACAGTA TAAATAGTTG	1800
	TAAGTTTGCA TGCATGATGG AAAAATAAAA ACCTGTATCT CTGTTAAAAA AAAAAAAAAA	1860
	AAAAAAAAA AAAAAAAAAA AAA	1883
55		

(2) INFORMATION FOR SEQ ID NO: 171:

60

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2100 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 171:

	TACTTTTAGA TTTACTGCCT TCAAAAAGTG CCTATCTGA GCAACATAAA CGTTATTCCT	60
10	TACATATGTA TGTACACACG GTACCCAGAG TCGTACTGTG GCAGCCTTCA AAAACATACC	120
	ATCAGAAAGA GTAGGTGCTG AGATAAGGNA ACTTTGCCAA ATGNAAGAAA GTCACTCACT	180
	TCCAATATCC CCTCTTCAAG CGGCTACCGT GRAASGGGCT GCAAACACAT TCCCTGAGCA	240
15	TCCCTTGCTG ATACAGCTTC TTTATATTTA TATCCTACTG GATGGTAGCA TATTGCTAAG	300
	GTTTCCTGTA CTCTGCTTCA AGGGAATGTA AGYTTTATGG CATTGAAACA TTTAGGAAAA	360
20	AAAAAGATGT TTAAGAGAAT TAATAGAGCC GTAGTCTGTA TTAGGATGTG TGTCATATGT	420
	GTGTTCTATA AACTAAGCAT CGGTGGGTTT AGAGTGTTAA AGTGTCAGCA CATTCTTCT	480
	CCTTTTGCTC CTCAGGCTAA CATGAGAGAA AATAGAAAAG TCTTGGCTGT GGGGATTGGA	540
25	AGCTCAGGGG GCCAAATGTC CTTGCCAGAT CCTTAGAGCA TFACTTTGAC TCCTAAAAAT	600
	AGTAGTGTAT GTTATTTGAT GGCTTTTGT TCCATAGTTC CATCACTGAC AAAACTGTCA	660
30	ATACTGTTGA TGGAGCAGCA GCATAGCCTA GAGTGATGCA TTCTTACCCA GAGGTGGCAA	720
	TAGGAGAGGG TCCATGTAAA TAGGACGAGG TAGACAGTGC ATGATTGTAG GAGAAGGGTT	780
	GAAGGGAGGA CATGATTCCA AAAAAGATCG TTCTCAATGT GTCGTCTGAC TCAACCAGCT	840
35	GGCAGATTAC ACTTGCCAAG TCGTTCCCTT TCCTTCTAAG TCAGTTGGCT CCATATTAC	900
	TTGAATATGC CTCTGTTTGG GCAAAGCAAG ATACCTCCAC TTAACCTTTA TCCAAGGAAG	960
40	CTCTTGGTGT CCTCTTGGTC ATAAAGTTGT CTCCTACCTA ACCCAGTTT ACCAAATGGA	1020
	AGTAAAAGGG GACAACTAT GGAAGATGGA CTCCATGCCA TTGCAGTCAG CCACCATTC	1080
	CTTTTCCATA TAAGGAGCCC CATTACATAA GCTACGGGTG AGGTTGGAAC AGCTATGTTT	1140
45	CATAATTTC AAGAGTGAC CACCCTGCTC TAGTCATCAT CATTGGATGA ATCCAGTTGA	1200
	CTCTTTGGCA AAAGGTGAT ACTTTTCACT AAAAATGCCT ACTCTTCCTG TTGATGTTCC	1260
50	TTTCTGTTT TTACCTTGTC CAATTCCAC ACTAGTCATT TTTTATTAT TTTAGAGGAT	1320
	CAGATTTTAG CGCTGAAAA TGAGTTCAAA AATTTCAGTG TAATGTCATA AGGATGTTGG	1380
	GATACAGAGA TTTTPTTTT CCTTGGAAC AAATGGACTG GGAAGAAACA CAGCATGGCT	1440
55	TTGCTCTGAG TTTCAATCTG ATGATTATGA CCATGGAAGA TAGTCTTATG TAAAGGTTAA	1500
	ATGGTGTTTA CAAGTGATA GATAAGGCGG AGATGGTGAG AAGCCGGGTT TTCTCTATGC	1560
60	TAAATGTGTC TACTAAGAGC AGCACTTCCT ACTAGCTAAG CACAATCATA GCCCCACCGT	1620

5 GATGAGCTGC TAGTCTGAAT AACATTCCTT GACTTAGGGA AAGGCACACA AAAACATATA 1680
AAGAATATGT CTATTTTCAT ATGTGTGATA CTGACAGAGC CATGGTATTC CTAAAATATA 1740
GGTTTCTCTT TTTTCTTGTA TTCTTAGCAA ATTGCATTTA TTCACTACAT TACAAACCAT 1800
CACTGATGTA TCCAAAATAG CACACATAGT TCAGTATGAA AATAAGAGAA TAAAATCTGT 1860
10 TATAAGCAAG TGATTTAGGT ATTTTCTTTT GTGTTTATGC ATTATCTGAC TATATTAAAA 1920
CCTGTTTTTC TATTTACCTT CTATCAGTTT TCTCTACCAA TTATGTTTTT TCAATGCTCT 1980
ATAAGAATGA ATATGGAAAT TATATTTCTT TTTTCTGTAA AAGAGTTGCA ACTACTTTAT 2040
15 TATATTTAGA AATCCAATAA ACTTCTTATT ACATTTAAAA AAAAAAAAAA AAAACTCGAA 2100

20

(2) INFORMATION FOR SEQ ID NO: 172:

(i) SEQUENCE CHARACTERISTICS:
25 (A) LENGTH: 1930 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 172:

30

CCTTTGANTG TGGTCCCGGG TGCNGATGG CAGCGCCTCC GCCGCGGCTC GTGGTTGTCC 60
CGCCATGGCA CTGTGCGGGG GGCTGCCCGG GGAGCTGGCT GAGGCGGTGG CCGGGGGCCG 120
35 GGTGCTGGTG GTGGGGGCGG GCGGCATCGG CTGCGAGCTC CTCAAGAAATC TCGTGCTCAC 180
CGGTTTCTCC CACATCGACC TGATTGATCT GGATACTATT GATGTAAGCA ACCTCAACAG 240
ACAGTTTTTG TTTCAAAAGA AACATGTTGG AAGATCAAAG GCACAGGTG CCAAGGAAAG 300
40 TGTACTGCAG TTTTACCCGA AAGCTAATAT CGTTGCCTAC CATGACAGCA TCATGAACCC 360
TGACTATAAT GTGGAAATTT TCCGACAGTT TATACTGGTT ATGAATGCCT TAGATAACAG 420
45 AGCTGCCCGA AACCATGTTA ATAGAATGTG CCTGCGAGCT GATGTTCCCT TTATTGAAAG 480
TGGAACAGCT GGGTATCTTG GACAAGTAAC TACTATCAAA AAGGGTGTGA CCGAGTGTTA 540
TGAGTGTCAT CCTAAGCCGA CCCAGAGAAC CTTTCCTGGC TGTACAATTC GTAACACACC 600
50 TTCAGAACCT ATACATTGCA TCGTTTGGGC AAAGTACTTG TTCAACCACT TGTTTGGGGA 660
AGAAGATGCT GATCAAGAAG TATCTCCTGA CAGAGCTGAC CCTGAAGCTG CCTGGGAACC 720
55 AACGGAAGCC GAAGCCAGAG CTAGAGCATC TAATGAAGAT GGTGACATTA AACGTATTTT 780
TACTAAGGAA TGGGCTAAAT CAACTGGATA TGATCCAGTT AAACCTTTTTA CCAAGCTTTT 840
TAAAGATGAC ATCAGGTATC TGTGACAAT GGACAACTA TGGCGGAAAA GGAAACCTCC 900
60

AGTTCCGTTG GACTGGGCTG AAGTACAAAG TCAAGGAGAA GAAACGAATG CATCAGATCA 960
ACAGAATGAA CCCCAGTTAG GCCTGAAAGA CCAGCAGGTT CTAGATGTAA AGAGCTATGC 1020
5 ACGTCTTTTT TCAAAGAGCA TCGAGACTTT GAGAGTTCAT TTAGCAGAAA AGGGGGATGG 1080
AGCTGAGCTC ATATGGGATA AGGATGACCC ATCTGCAATG GATTTTGTCA CCTCTGCTGC 1140
10 AACCTCAGG ATGCATATTT TCAGTATGAA TATGAAGAGT AGATTTGATA TCAAATCAAT 1200
GGCAGGGAAC ATTATTCCTG CTATTGCTAC TACTAATGCA GTAATTGCTG GGTGTAGTAGT 1260
ATTGGAAGGA TTGAAGATTT TATCAGGAAA AATAGACCAG TGCAGAACAA TTTMTTTGAA 1320
15 TAAACAACCA AACCCAAGAA AGAAGCTTCT TGTGCCTTGT GCACTGGATC CTCCCAACCC 1380
CAATGTGTAT GTATGTGCCA GCAAGCCAGA GGTGACTGTG CGGCTGAATG TCCATAAAGT 1440
GACTGTTCTC ACCTTACAAG ACAAGATAGT GAAAGAAAAA TTTGCTATGG TAGCACCAGA 1500
20 TGTCCAAATT GAAGATGGGA AAGGAACAAT CCTAATATCT TCCGAAGAGG GAGAGACGGA 1560
AGCTAATAAT CACAAGAAGT TGTGAGAATT TGGAAATTAGA AATGGCAGCC GGCTTCAAGC 1620
25 AGATGACTTC CTCCAGGACT ATACTTTATT GATCAACATC CTTCATAGTG AAGACCTAGG 1680
AAAGGACGTT GAATTTGAAG TTGTTGGTGA TGCCCCGGA AAAGTGGGGS CCAAACAAGC 1740
TGAAGATGCT GCCAAAAGCA TAACCAATGG GCAGTGATGA TGGGAGCTTC AGCCCTCCAC 1800
30 CTYCACAGCT TCAAGGAGGC AAGATGGACG TYTCYCATAG TTGATYCGGR TGAAGAAGRT 1860
TCTCCAATAA TTGCCCGACG TTCATTGAAG GAAGGAGGAG GAGGCCCGCC AAGAGGGGAA 1920
35 TTTAGGNITG 1930

40 (2) INFORMATION FOR SEQ ID NO: 173:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1509 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 173:

50 GGCCCTGGCC TCTGGGCTGA GGCTTGCTAG GGAATCGGGG TGGCTCTAAG GGGCAGGGAT 60
AGGGCTGGGG AGCGCCGGCC TGTGGCCCTG ACCAGCCCCT TCTCGTGCRC GTTCCACCCC 120
GATGCAGGTG GTCACGTGCT TGACGCGGGA CAGCTACCTG ACGCACTGCT TCCTCCAGCA 180
55 CCTCATGGTC GTGCTGTCTT CTCTGGAACG CACGCCCTCG CCGGAGCCTG TTGACAAGGA 240
CTTCTACTCC GAGTTTGGGA ACAAGACCAC AGGGAAGATG GAGAACTACG AGCTGATCCA 300
60 CTCTAGTCGC GTCAAGTTTA CCTACCCAG TGAGGAGGAG ATTGGGGACC TGACGTTTAC 360

	TGTGGCCCAA AAGATGGCTG AGCCAGAGAA GGCCCCAGCC CTCAGCATCC TGCTGTACGT	420
5	GCAGGCCCTTC CAGGTGGGCA TGCCACCCCC TGGGTGCTGC AGGGGCCCCC TCGGCCCCAA	480
	GACACTCCTG CTCACCAGCT CCGAGATCTT CCTCCTGGAT GAGGACTGTG TCCACTACCC	540
	ACTGCCCGAG TTTGCCAAAG AGCCGCCGCA GAGAGACAGG TACCGGCTGG ACGATGGCCG	600
10	CCGCGTCCGG GACCTGGACC GAGTGCTCAT GGGCTACCAG ACCTACCCGC AGCCCTCACC	660
	CTCGTCTTCG ATGACGTGCA AGGTCATGAC CTCATGGGCA GTGTCAACCCT GGACCACTTT	720
15	GGGGAGGTGC CAGGTGGCCC GGCTAGAGCC AGCCAGGGCC GTGAAGTCCA GTGGCAGGTG	780
	TTTGTCCCCA GTGCTGAGAG CAGAGAGAAG CTCATCTCGC TGTGGCTCG CCAGTGGGAG	840
	GCCCTGTGTG GCCGTGAGCT GCCTGTCGAG CTCACCGGT AGCCAGGGC ACAGCCAGCC	900
20	TGTGCTGTCC AGCCTGACGC CTACTGGGC AGGGCAGCAG GCTTTTGTGT TCTCTAAAAA	960
	TGTTTTATCC TCCCTTTGGT ACCTTAATTT GACTGTCTTC GCAGAGAATG TGAACATGTG	1020
25	TGTGTGTGT GTTAATTCTT TCTCATGTTG GGAGTGAGAA TGCCGGGCC CTCAGGGCTG	1080
	TCGGTGTGCT GTCAGCCTCC CACAGGTGGT ACAGCCGTGC ACACCAAGTGT CGTGTCTGCT	1140
	GTTGTGGGAC CGTTGTAAAC ACGTGACACT GTGGGTCTGA CTTTCTCTTC TACACGTCTT	1200
30	TTCTGAAGT GTCGAGTCCA GTCCTTTGTT GCTGTTGCTG TTGCTGTTGC TGTGCTGTT	1260
	GGCATCTTGC TGCTAATCCT GAGGCTGGTA GCAGAATGCA CATTGGAAGC TCCCACCCCA	1320
35	TATGTCTCTT CAAAGTGGAG GTCTCCCCTG ATCCAGACAA GTGGGAGAGC CCGTGGGGGC	1380
	AGGGGACCTG GAGCTGCCAG CACCAAGCGT GATTCTGCT GCCTGTATTC TCTATPCCAA	1440
	TAAAGCAGAG TTTGACACCG TCAAAAAAAA AAAAAAAAAA AAAAAAAAAA ATTNCTGCGG	1500
40	CCTCAAGGG	1509

45 (2) INFORMATION FOR SEQ ID NO: 174:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 3173 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 174:

55	TCGACCCAS GGTCCGTGC TTTCCACAG AAGTTAGAC CCTGAAAGAG ATGGCTCAGC	60
	ACCACCTATG GATCTTGCTC CTTGCCTGC AAACCTGGCC GGAAGCAGCT GGAAAAGACT	120
60	CAGAAATCTT CACAGTGAAT GGGATTCTGG GAGAGTCAGT CACTTCCCT GTAAATATCC	180

	AAGAACCACG GCAAGTTAAA ATCATTGCTT GGACTTCTAA AACATCTGTT GCTTATGTAA	240
	CACCAGGAGA CTCAGAAACA GCACCCGTAG TTAGTGTGAC CCACAGAAAT TATTATGAAC	300
5	GGATACATGC CTTAGGTCCG AACTACAATC TGGTCATTAG CGATCTGAGG ATGGAAGACG	360
	CAGGAGACTA CAAAGCAGAC ATAAATACAC AGGCTGATCC CTACACCACC ACCAAGCGCT	420
10	ACAACCTGCA AATCTATCGT CGGCTTGGGA AACCAAAAAT TACACAGAGT TTAATGGCAT	480
	CTGTGAACAG CACCTGTAAT GTCACACTGA CATGCTCTGT AGAGAAAGAA GAAAAGAATG	540
	TGACATACAA TTGGAGTCCC CTGGGAGAAG AGGGTAATGT CCTTCAAATC TTCCAGACTC	600
15	CTGAGGACCA AGAGCTGACT TACACGTGTA CAGCCAGAA CCCTGTCAGC AACAAITCTG	660
	ACTCCATCTC TGCCCGGCAG CTCTGTGAG ACATCGCAAT GGGCTTCCGT ACTCACCACA	720
20	CCGGGTGCT GAGCGTCTG GCTATGTTCT TTCTGCTTGT TCTCATTTCTG TCTTCAGTGT	780
	TTTTGTTCG TTTGTTCAG AGAAGACAAG ATGCTGCCTC AAAGAAAACC ATATACACAT	840
	ATATCATGGC TTCAAGGAAC ACCCAGCCAG CAGAGTCCAG AATCTATGAT GAAATCCTGC	900
25	AGTCCAAGGT GCTTCCCTCC AAGGAAGAGC CAGTGAACAC AGTTTATTCC GAAGTGCAGT	960
	TTGCTGATAA GATGGGGAAA GCCAGCACAC AGGACAGTAA ACCTCCTGGG ACTTCAAGCT	1020
30	ATGAAATTGT GATCTAGGCT GCTGGGCTGA ATTCTCCCTC TGGAACTGA GTTACAACCA	1080
	CCAATACTGG CAGGTCCCTT GGATCCAGAT CTCTCTGCC CAACTCTTAC TGGGAGATTG	1140
	CAAACGCCA CATCTCAGCC TGTAAAGCAA GCAGGAAACC TTCTGCTGGG CATAGCTTGT	1200
35	GCCTAAATGG ACAAATGGAT GCATACCTT CCTGAAATGA CTCCCTTCTG AATGAATGAC	1260
	AAAGCAGGTT ACCTAGTATA GTTTTCCCAA ACTTCTTCCC ATCATAGCAC ATGTAGAAAA	1320
40	TAATAITTTT ATGGCACACT GGGATAAACA AGCAAGATTG CTCACTTCTG GAAGCTGCAT	1380
	ATGACTAGAG GCCTCTGTG ACTGGAGGTA ACAACCCTGC CCAGTAACTG TGGGAGAAGG	1440
	GGATCAATAT TTTGCACACC TGTAAATAGG CATGGCACAC CAGCCAAGAT GCTCTGCTCA	1500
45	CAGTCAGTAT GTGTGAAGAT CCCTGGTGCG TGGCCTTCAC CACGCATCTT GAGCAAATTA	1560
	GGAAAATGTA CCCTTCGCTT GAGGCAGATG CAGCCCTTCC CCCGAGTGCA TGGCTTGGAG	1620
50	AGCAGAATGT GGGCTGCATA TAAGCACACT CATCCCTTTG TCTGGGAATC TTTGTGCAGG	1680
	GCATAACAGG CTTAGTAAGT CCAAACACAG ATGACAGTGC TGTGTGGGTC TCTGTCAGAG	1740
	TTGTGGCTCT CAGCCATGTA GACACACTCT CCAAATGGAG TGTGGAAAA TGTTCITTTCT	1800
55	GCAGGGTCTA GAGACTGCTG GGACACTTTT CTTGGAGTGC TACTTCAGAA GCCTTATAGG	1860
	ATTTTCTTTT TGGCCAAGAT TTCTTCTGT ATCACTCCAA GCAGCCTCAG CAGAAGAAGC	1920
60	AGCCATGCCC AGTATTCCCA CTCTCCAAAA GGAACGTACC AGCTTATATT TCTCACACTT	1980

	CTGGGGA	ACT	GGTATAATC	CAACCATCAA	AATAGAAGAC	CTTGCAAGAA	GCAGAGTCAT	2040
	TCTCCAGAAG	GA	ACTTGGGA	GATGATGGTG	CAGATGATGA	AACTGGGTTC	ATCCCAGTTC	2100
5	CAAAGACTCA	GAGAACTAGA	GTTTAAGCTG	AGGCAGAGTG	CCGCCACCCT	GGCATGCCCC		2160
	ACAAACAGAT	CACCAGCCAG	CTTACACAGG	CATTAACTCT	CCTCAATGAG	GAAGAATCAT		2220
	TCACA	ACTGA	GCAAGACATT	CATATGATCA	TTTAAGGAAG	TGTTTCCCTT	ATGTGTTAGC	2280
10	AAGTATAATC	GGCTAACTCC	TAAATCCCAA	TGAATAGTCC	TAGGCTGGAC	AGCAATGGGC		2340
	TGCAATTAGG	CAGATAAAGA	CATCAGTCCC	AGTAAATGAA	TCCATAGACT	CATCTAGCAC		2400
15	CAACTACCAT	TAGCACTATG	TTAGGAGCTG	CAAGCCCCA	AAGTAGAAGA	TGTGCATAAT		2460
	GTCTGCTCTT	GTGTAGCTCA	GGAGACAATT	CCAGCACAGA	CACTACAGTT	AACGCTGAAC		2520
	TGCAGCTGCA	AGTAATAGCA	TGAACAGTCA	GAAAAATACC	TTATGAGGGG	GCAGGGCTGA		2580
20	AGCTGGGCCT	TGAAGGATGG	ATGAAATTTG	GATAGAGAAT	GAGGAAGACA	GAGGGCCTCC		2640
	AAGTGAGAGA	AGCATGAAAA	ATGAGCAGGG	GCCTGGATCA	GTGGGGTGTA	TTCAGAGCAC		2700
25	CTCTCCAGAT	GCACCATGCA	TGCTCACAGT	CCCTTGCCTA	TGTGTGGCAG	AGTGTCCAG		2760
	CCAGATGTGT	GCCCCCACCC	CATGTCCATT	TACATGTCCT	TCAATGCCCA	CCTCAAAAGG		2820
	TACCTCTTCT	GTAAAGCTTT	CCCTGGTATC	AGGAATCAAA	ATTAATCAGG	GATCTTTTCA		2880
30	CACTGCTGTT	TTTTCTCTTT	TGGTCCTTCT	ATCACTAAAA	CTCATCTCAT	TCAGCCTTAC		2940
	AGCATAACTA	ATTATTGTGT	TTCTCTACTA	CATTGTACAT	GTGGGAATTA	CAGATAAACG		3000
35	GAAGCCKGCT	GGGGTGGTGG	CTCAGCCTG	TAATCCCAAC	ACTTTGGGAG	GCCAAGGCAG		3060
	GCGGATCACC	TGAGGTCAGG	ARTTCGAGAT	TARTCTGGCC	AACATGGTGA	AACCCCATMT		3120
40	NTACTAAAAA	TACGAAATTA	GCCAGGTGTG	GTGGCACACA	TCTGTAGTCC	CAG		3173

(2) INFORMATION FOR SEQ ID NO: 175:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 991 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 175:

	AAATT	CGGCA	CAGCTGAGAG	GAGACACAAG	GAGCAGCCCG	CAAGCACCAA	GTGAGAGGCA	60
55	TGAAGTTACA	GTGTGTTTCC	CTTTGGCTCC	TGGGTACAAT	ACTGATATTG	TGCTCAGTAG		120
	ACAACCACGG	TCTCAGGAGA	TGTCTGATTT	CCACAGACAT	GCACCATATA	GAAGAGAGTT		180
60	TCCAAGAAAT	CAAAAGAGCC	ATCCAAGCTA	AGGACACCTT	CCCAAATGTC	ACTATCCTGT		240

CCACATTTGA GACTCTGCAG ATCATTAAAGC CCTTAGATGT GTGCTGCGTG ACCAAGAACC 300
5 TCCTGGCGTT CTACGTGGAC AGGGTGTCA AGGATCATCA GGAGCCAAAC CCCAAATCT 360
TGAGAAAAAT CAGCAGCATT GCCAACTCTT TCCTCTACAT GCAGAAAACT CTGCGGCAAT 420
GTCAGGAACA GAGGCAGTGT CACTGCAGGC AGGAAGCCAC CAATGCCACC AGAGTCATCC 480
10 ATGACAACTA TGATCAGCTG GAGGTCCACG CTGCTGCCAT TAAATCCCTG GGAGAGCTCG 540
ACGTCTTTCT AGCCTGGATT AATAAGAATC ATGAAGTAAT GTCCTCAGCT TGATGACAAG 600
GAACCTGTAT AGTGATCCAG GGATGAACAC CCCCTGTGCG GTTTACTGTG GGAGACAGCC 660
15 CACCTTGAAG GGAAGGAGA TGGGAAGGC CCCTTGCAGC TGAAAGTCCC ACTGGCTGGC 720
CTCAGGCTGT CTTATTCCGC TTGAAAATAG CAAAAAGTC TACTGTGGTA TTTGTAATAA 780
20 ACTCTATCTG CTGAAAGGGC CTGCAGGCCA TCCTGGGAGT AAAGGGCTGC CTTCCCATCT 840
AATTTATTGT GAAGTCATAT AGTCCATGTC TGTGATGTGA GCCAAGTGAT ATCCTGTAGT 900
ACACATTGTA CTGAGTGGTT TTTCTGAATA AATTCCATAT TTTACCTAAA AAAAAAAAAA 960
25 AAAAATCGA GGGGGGGCCC GTACCCAATT T 991

30

(2) INFORMATION FOR SEQ ID NO: 176:

(i) SEQUENCE CHARACTERISTICS:
35 (A) LENGTH: 1290 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 176:

ACAGCCCTCT TCGGAGCCTG AGCCCGGCTC TCCTCACTCA CCTCAACCCC CAGGCGGGCC 60
CTCCACAGGG CCCCTCTCCT GCCTGGACGG CTCTGCTGGT CTCCCCGTCC CCTGGAGAAG 120
45 AACAAAGCCA TGGGTGGGCC CCTGCTGCTG CCCCTRCTGC YCCTGCTGCW GCCGCCAGCA 180
TTTCTGCAGC CTRGTGGCTC CACAGGATCT GGTCCAAGCT ACCTTTATGG GGTCACTCAA 240
CCAAAACACC TCTCAGCCTC CATGGGTGGC TCTGTGAAA TCCCTTCTC CTTCTATTAC 300
50 CCCTGGGAGT TAGCCAYAGY TCCCRACGTG AGAATATCCT GGAGACGGGG CCACTTCCAC 360
GGGCAGTCCT TCTACAGCAC AAGCCGCCT TCCATTCACT AGGATTATGT GAACCGGCTC 420
55 TTTCTGAACT GGACAGAGGG TCAGGAGAGC GGCTTCCTCA GGATCTCAA CCTGCGGAAG 480
GAGGACCAGT CTGTGTATTT CTGCCGAGTC GAGCTGGACA CCCGGAGATC AGGGAGGCAG 540
CAGTTGCAGT CCATCAAGGG GACCAAACTC ACCATCACCC AGGCTGTCAC AACCACCACC 600
60

ACCTGGAGGC CCAGCAGCAC AACCAACATA GCCGGCCTCA GGGTCACAGA AAGCAAAGGG 660
CACTCAGAAT CATGGCACCT AAGTCTGGAC ACTGCCATCA GGGTTGCATT GGCTGTCGCT 720
5 GTGCTCAAAA CTGTCAATTT GGGACTGCTG TGCCTCCTCC TCTGTGGTGG AGGAGAAGGA 780
AAGGTAGCAG GCGGCCAAGC AGTGAATTCT GACCAACAGA GTGTGGGGAG AAGGGATGTG 840
TATTAGCCCC GGAGGACGTG ATGTGAGACC CGCTTGTGAG TCCTCCACAC TCGTTCCCCA 900
10 TTGGCAAGAT ACATGGAGAG CACCCTGAGG ACCTTTAAAA GGCAAAGCCG CAAGGCAGAA 960
GGAGGCTGGG TCCCTGAATC ACCGACTGGA GGAGAGTTAC CTACAAGAGC CTTTCATCCAG 1020
15 GAGCATCCAC ACTGCAATGA TATAGGAATG AGGTCTGAAC TCCACTGAAT TAAACCACTG 1080
GCATTGGGG GCTGTTYATT ATAGCAGTGC AAAGAGTTCC TTTATCCTCC CCAAGGATGG 1140
AAAATACAAT TTATTTTGCT TACCATACAC CCTTTTCTC CTCGTCCACA TTTTCCAATC 1200
20 TGTATGGTGG CTGTCTCTA TGGCAGAAGG TTTTGGGGAA TAAATAGCGT GANATGNTNC 1260
TGACTNAAAA AAAAAAAAAA AAAAATCGA 1290

25

(2) INFORMATION FOR SEQ ID NO: 177:

30

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2290 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 177:

TGGGGCCCCCT TTTGGATGCT CTGGGTGTTT TTGCCAAGAG TTACAGGATG TCAAGTGTGG 60
40 GGAGCTCAGC ACCCTTGCTG TGGACCAGTG AAGGCTGTTC CAGACCAGGT GCTTCCAGAC 120
ATTTCCAGGC TCCAGGAGAG AGGCTGGGAG CCCCCACAGA AAGCACAGGA AAATGCAAAA 180
AAAAACAGT CTTTTTTTTT TTTTGTCTT TTATTATGAA AACAAAACAA ATGCCCCAGG 240
45 AGAAGGGTCC ATGATTACCA GAAACATCAA AGAGTACTTT CTACCATTTT TATCTGTGTG 300
TGTTGAGGCC AGCATTGCAA TAAACAAGCT AAATACTTA CATTTGACTC ATTTTCAGTA 360
50 ACTGACATTT ACAGGAATAT ACTAGAAACG GCACTAAAAA GTTTAAGAAA AGTTACGGTA 420
AACTTGCATG CACATCATAC AGAAAAGTAA CATTTTAAAT ATAAAAAGA AAAACTTCCT 480
GGAAGCATTA TGCCAGTATT AAGGAACAGT GCTACTCTGG ATGTGACAAA TTCTGTATGT 540
55 GGGTGTACT CTTTCCCAA AGACTGTGAG AGGCGTGAGT GCTGCAAAAG AACACAACA 600
AAAACAACA CACAAAAAA TGTGTCTTAC AGTTTGTAAG CAAGATGACA CTGCCCCACA 660
60 CAAAGAGGGG TCTGGAGTTC AGTTACGCC CGAAGCCTGC CCCCTCGGCC TCCAGGGGTC 720

	ATTTCAGAGTG TTCTCAAATC CAATTCGGAC ACACGACTTG TCACTACTCC TCTCCCCTTG	780
5	AAAAAAGCAT GTTAGAAGCT GCCCTACAGG TCTCAGCAGT GGGACAATCT AATTGAATCA	840
	CCGCAGCCTT CTAATACAGA AGAAACGGAC GTGACTGTCA CCCTCAGCCC GCCAGCAAGG	900
	GCGCTGAGGA AGTCATTAAT CCTTCGAAAC TCTGAAAAGA AACCAGTGTG GAAGTCTGGA	960
10	CAGAAAGCCT TAAAAAAGTG ACAGCACCAA TGCAGCTGCT CAGTGTACCC NCCGTGGGCT	1020
	GTCAGGGTCA GTGGCTTCTT TCTAGATGAA AGGAGCAGAG GCGAGCCGAC GCCACCGTCA	1080
15	CAGAGAACCA GCCGAGAAGG AAAGGCCCCA CGATGCTCCC TGTCGCTGCG CCCACAGCC	1140
	GGCGCTTCCC CCGACGGCTC ACACAGGCAG CACCTCACTG CCCTGTGGCT GGAGGGGCAT	1200
	TGCAAGGAGC GCGCCCGAGC CCCAGGCACC CCGGCTTAG GGTGTACGTA TCACCCAGCC	1260
20	CTGTGCTGGC AGCACGTTAC CAACCAGCCT GCGTGAAGAC CTGTCAACTG TCGTGTGTGA	1320
	ATTCCTTAAA TTCGGTTTAA ATAGTCCATT AAAGATCTGT TTAGAAAATA CCTTTGAAAA	1380
25	CGAGGGTAAC TTTAAAAAAT GGAAACTTTC AAATCCATTT ATATTTTAT TATAAACAAA	1440
	ACTTAATTAA AAGTTTAAAC AACTGGCTGA AAACCTACCA AGTGTACAGC TCACCAGCAA	1500
	TTTAAAAAAT GATAATTTAC CAGCATCTCC TCATCAGAGT TCCCTCTCCA GTAAGGGTAT	1560
30	ACCTACATCT GTAAGGGTCA GTGGACTCTG AATCAATTTT ATGGTTGTIT TAAAATCACC	1620
	GTGTATTAGG ATACTAATGA TAGTCCCTAT ATCCATCCAG AAATGCTGGC AGAAAGCACT	1680
35	GGCCACCATA CAGGACAGAC CACACCACAG CTCCATACCC AGCGTCTGCC TGGAGGCTCC	1740
	CCCACGCTGA GGTCCGGGAG AATGCCTGGT TTCAGTCATT TCCGGACTAA CTGTGACAAC	1800
	GCGTGAGCAG GGAGCACCGT GCGAGTCTCC GGGAGGGAAT CCTCCTGGGG CCCAGAGACT	1860
40	CCTCCACCCC TGGGGAGGGC AGACAGGCTC GGGARGGCCT GGCCAGGCCA CTGGAGGCTG	1920
	GCAGGGAGCA GGCATGTCCA CCCGCAAGCC TGGGAGGCTA ACTCTGGCAT TCCTGGCCGG	1980
45	AGCCGCCATG CTCATTGGTG GGCCAGTTTG GGACATCCCC GTACTCAAAG ACCATATGGC	2040
	AGCCTCTGGG AAAACAAAAC CAAAACATCA CCTTCTATTA AACTCTGTAT ATTATTATTT	2100
	TTTACAATAG AAAGTTAAAA ATCAAGACTT AGATTTACTA TACATTTTTT CTCTCAGATT	2160
50	ACAAAGTTTA TATTATATAA CTGGGGTCC CTAAATGAT TTCTTTTAAA ACAGTCTTAA	2220
	AGAGACCAGA AGTGAATACA AAAGAACTAA ACAAATAAAA AAATTAGAAT GTGCTGTAGC	2280
55	TGAAAGCTGT	2290

(2) INFORMATION FOR SEQ ID NO: 178:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 549 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 178:

10 GGCACGAGCC ATGCTTGGCC TCTCCTTGAT TCTTACAGTC ACTTTGTTGG CTGTTTCTGA 60
CTCAGCAGCT ACCTGCATTG TGGCCAAAGG ATGACCTATT CCTTCTCAGG AGGGCAAAAA 120
TGTGGAATAG TGTCTGTCCA TGCCTCTCCT CATGGGCTAC CACCTCTGCC ACCGTGGTTA 180
15 ATCAGTAACA ACCAGGAGAG AAGCTGCTGG AACTGACCTC TGGGAACTCC CTGGGATGGT 240
TTGGTGCAGG AATGTAGTAG GCATACACGT GGTTCGCTGG ATCTGGGCC TCCTGATGTG 300
AGTAGAGAGG TAAAAGGCCA CCATCTCCTT GACCTCTGGG GAACTCATCC ACAAAGAAGA 360
20 TGTTCCTCAAG ATGCTTCTGA AGATTGCCTA AAAATAGCCG GTTCCACCC CCGTGAATGC 420
ATCCATCTTA GAATGCTCCT TCACCAGGAC CAGAGAACTG ATTTACAGAA GTGACATGAA 480
25 AACATTCAT CCCAGAATTT GCAGTAGCTC AAATTAAGTT TCTAGCTATT AAAAAGAAAA 540
AAAAA 549

30

(2) INFORMATION FOR SEQ ID NO: 179:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1509 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

35

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 179:

GGCACGAGGG CTCATTCAAT CCGCGCGGG CCTGCCAGAC ACCTGCGCCC TTCTGCAGCC 60
45 GCGCGCGCA TCCGCCCGG CAGCCCCAG CATGTCGGG CCAGACGTGC AGACGCCGTC 120
CGCCATCCAG ATCTGCCGA TCATCGGCC AGATGATGCC AACGTGGCG GCAATGTCCA 180
CGGGGGACC ATCTGAAGA TGATCGAGGA GGCAGGCGCC ATCATCAGCA CCCGGCATTG 240
50 CAACAGCCAG AACGGGAGC GCTGTGTGG CGCCCTGGCT CGTGTGAGC GCACCGACTT 300
CCTGTCTCCC ATGTGCATCG GTGAGGTGGC GCATGTCAGC GCGGAGATCA CCTACACCTC 360
CAAGCACTCT GTGGAGTGC AGGTCAACGT GATGTCCGAA AACATCCTCA CAGGTGCCAA 420
55 AAAGCTGACC AATAAGGCCA CCCTGTGTA TGTGCCCTG TCGCTGAAGA ATGTGGACAA 480
GGTCTCGAG GTGCCTCTG TTGTGTATTC CCGGCAGAG CAGGAGGAG AGGGCCGAA 540
60 GCGGTATGAA GCCCAGAAGC TGGAGCGCAT GGAGACCAAG TGGAGGAACG GGGACATCGT 600

CCAGCCAGTC CTCAACCCAG AGCCGAACAC TGTGAGTAC AGCCAGTCCA GCTTGATCCA 660
CCTGGTGGGG CCTTCAGACT GCACCCCTGCA CGGCTTTGTG CACGAGGGTG TGACCATGAA 720
5 GCTCATGGAT GAGGTCGCCG GGATCGTGGC TGCACGCCAC TGCAAGACCA ACATCGTCAC 780
AGCTTCCGTG GACCCATTA ATTTTCATGA CAAGATCAGA AAAGGCTGCG TCATCACCAT 840
10 CTCGGGACGC ATGACCTTCA CGAGCAATAA GTCCATGGAG ATCGAGGTGT TGGTGGACGC 900
CGACCCGTGT GTGGACAGCT CTCAGAAGCG CTACCGGGCC GCCAGTGCCT TCTTCACCTA 960
CGTGTGCTG AGCCAGGAAG GCAGGTCGCT GCCTGTGCCC CAGCTGGTGC CCGAGACCGA 1020
15 GGACGAGAAG AAGCGCTTTG AGGAAGGCAA AGGGCGGTAC CTGCAGATGA AGGCGAAGCR 1080
ACAGGGCCAC GCGGASCYTC AGCCCTAGAC TCCCTCCTCC TGCCACTGGT GCCTCGAGTA 1140
20 GCCATGGCAA CGGGCCAGT GTCCAGTCAC TTAGAAGTTC CCCCTTGGC CAAAAACCCA 1200
ATTCACATTG AGAGCTGGTG TTGTCTGAAG TTTCGTATC ACAGTGTAA CCTGTACTCT 1260
CTCCTGCAA CCTACACACC AAAGCTTTAT TTATATCATT CCAGTATCAA TGCTACACAG 1320
25 TGTGTGCCG AGCGCCGGA GCGTGGGC AGAAACCTC GGAATGCTT CCGAGCACGC 1380
TGTAGGTAT GGGAAGAACC CAGCACCCT AATAAGCTG CTGCTGGCT GGAAAAAAA 1440
30 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1500
AGAAAAAAN 1509

35

(2) INFORMATION FOR SEQ ID NO: 180:

(i) SEQUENCE CHARACTERISTICS:
40 (A) LENGTH: 1316 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 180:

AGCTGTATCA TAGGAAAGAT GGCCACACCG GCGGTACCAG TAAGTGCTCC TCCGGCCACG 60
CCAACCCAG TCCCGCGGC GGGCCAGCC TCAGTTCCAG CGCCAACGCC AGCACC GGCT 120
50 GCGGCTCCG TTCCCGCTGC GGCTCCAGCC TGCATCTCA GACCTGCGG CAGCAGCGC 180
TGCAACTGCG GCTCCTGGCC AGACCCCGC CTCAGCGCAA NTCCAGCGCA GACCCAGCG 240
55 CCCGCTCTG CTGGTCTGC TCTTCCAGG CCTTCCCG GCGGCGCGT GGTGAGCTG 300
CACCCAGTCA TTTTGGCTC CATGTGGAC AGCTACGAGA GACGCAACGA GGTGCTGCC 360
CGAGTTATCG GGACCCGTG GGAACGTG GACAAACACT CAGTGGAGGT CACCAATTGC 420
60

TTTTCAGTGC CGCACAAATGA GTCAGAAGAT GAAGTGGCTG TTGACATGGA ATTTGCTAAG 480
AATATGTATG AACTGCATAA AAAAGTTTCT CCAAATGAGC TCATCCTGGG CTGGTACGCT 540
5 ACGGGCCATG ACATCACAGA GCACTCTGTG CTGNATCCAT GAGTACTACA GCCGAGAGGC 600
CCCCAACCCC ATCCACCTCA CTGTGGACAC AAGTCTCCAG AACGGCCGCA TGAGCATCAA 660
AGCCTACGTC AGCACTTTAA TGGGAGTCCC TGGGAGGACC ATGGGAGTGA TGTTCAGGCC 720
10 TCTGACAGTG AAATACGCGT ACTACGACAC TGAACGCATC GGAGTTGACC TGATCATGAA 780
GACCTGCTTT AGCCCCAACA GAGTGATTGG ACTCTCAAGT GACTTGCAGC AAGTAGGAGG 840
15 GGCATCAGCT CGCATCCAGG ATGCCCTGAG TACAGTGTG CAATATGCAG AGGATGTACT 900
GTCTGAAAG GTGTCAGCTG ACAATACTGT GGGCCGCTTC CTGATGAGCC TGGTTAACCA 960
AGTACCGAAA ATAGTTCCCG ATGACTTTGA GACCATGCTC AACAGCAACA TCAATGACCT 1020
20 TTTGATGGTG ACCTACCTGG CCAACCTCAC ACAGTCACAG ATTGCACTCA ATGAAAACT 1080
TGTAACCTG TGAATGGACC CCAAGCAGTA CACTTGCTGG TCTAGGTATT AACCCAGGA 1140
25 CTCAGAAGTG AAGGAGAAAT GGGTTTTTTG TGGTCTTGAG TCACACTGAG ATAGTCAGTT 1200
GTGTGTGACT CTAATAAACG GAGCCTACCT TTTGTAAATT AAAAAAAAAA AAAAAAACCN 1260
SGRGGGGGGG CCCGGTCCCA TTSSCCCTTT NGTAATTCGT NITACAATCC CCNGGC 1316
30

(2) INFORMATION FOR SEQ ID NO: 181:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 777 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 181:

GGCATGWKCA GACATGACTT CTATTGCCAG GCTGGTCAAG TGGCAGGGTC ATGAGGGAGA 60
45 CATCGATAAG GGTGCTCCTT ATGCTCCCTG CTCTGGAATC CACCAGCGGG CTATCTGCGT 120
TTATGGGGCT GGGGACTAGA ATTGGATGCT TCAAAACCAT CACCTGTTGG CCAACAAGTT 180
50 TGACCCAAAG GTAGATGATA ATGCTCTTCA GTGCTTAGAA GAATACCTAC GTTATAAGGG 240
CCATTCTATT GGGACCTGAA CTTTGAAGAC CACAMTATG AAGAGGCGTT GCTTACCYGT 300
TGGGGGCCAA GAGGCATGTT ACCAAACATG GYYCARGAAM YTTGGYKGGG AMCARKKKKG 360
55 GKKGGGARRM CMRGGGYTTG SCAAWTTCCK KGGCMWCCYT TTAGGGTAAR RRRGGCKGTW 420
ATTAGATTGT GGGTAAAGTA GGATCTTTTG CCCTTGCAA TTTGCTGCCT GGGTGAATGY 480
60 TGCTTGTTCC TTCTCMACCC CTAACCCTAG TAGTTCCTCC ACTAACTTTC TCACTAAGTG 540

AGAATGAGAA CTGCTGTGAT AGGAGAGTG AAGGAGGGAT ATGTGGTAGA GCACTTGATT 600
TCAGTTGAAT GCCTGCTGGT AGCTTTTCCA TTCTGTGGAG CTGCCGTTCC TAATAATTCC 660
5 AGGTTTGGTA GCGTGGAGGA GAACTTTGAT GGAAAGAGAA CCTTCCCTTC TGTACTGTTA 720
ACTTAAAAAT AAATAGCTCC TGATTCAAAG TAAAAA AAAA AAAA 777
10

(2) INFORMATION FOR SEQ ID NO: 182:

15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 791 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 182:

GGCACAGATA ACTATGTACA TGTATTCCTT AAATGTTTTT TTAAGTTTTA TATTCTTGCC 60
25 ACTGGTCTTC AAATGTGTAC ATGTGTGCCA GGGAGCAAAT GCCTTCTTGT TTCTGAAATT 120
GGTCTTTTAG ACTGTTCTTT TTCCCATCT TCTCACCTCC TGCCCCCTCT TCAGGGTACT 180
TCCGTGGCCA GAACCCCTCC AGGTCAGAGG CAGAAGAGAA GCCTCATGGG TCACAGCAGC 240
30 AGATGTGGGC TGGAGATCTA TTCATTGGT TTGGCTTGA ATTTCTGRA TGGTTTACTT 300
GATCYTGGGA AAGANATATC TTGCCAGGAA AAATGATAGN CCTTGACAAT GTTGAATGAT 360
35 CCTGCACCAC CTTGAAAGAC ATTTCTAATA TGGTTTGTC GGCAAAGTGG TTAGTAGTCA 420
TTTGTGGCCT GAGGTAGAAG TCCTCAGAAA TCAGCAGACT TCACTGATAA AATGCTGACT 480
TGCCCCCTGA CTGGGCTCTG TGAGAGTGGC CTCTGCACT GTGCACAGTA GGTGTGAACA 540
40 CACCACACCT ACAGGGACCA CGTGGTGGC TGTGGACTAG CGGCCAAGCT CCCTGCAGGC 600
CCACTAATAG AATTCAGCTT TTAGCATGGG CTGTTTCATA CTGTTCTGAT GAAACTGATT 660
45 TGGTTCTTT CCTCCATACC CCTTCTGCAT TTCAGTGTTC TTGTTTAGTT TTCCTGGTTT 720
TTAATTATAA CTACAAAATA AAATCTTTAG GCTATTCACC TTAGCTTAGT AAAAAAAAAA 780
AAAAAAACT C 791
50

(2) INFORMATION FOR SEQ ID NO: 183:

55 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1405 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
60 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 183:

5 AAATGTGATTA ACAGCTTGAA AGAAGGCTCT GGTTTTGAAG GCCTAGATAG CAGCACTGCC 60
AGTAGCATGG AGCTGGAAGA ACTTCGGCAT GAGAAAGAGA TGCAGAGGGA GGAAATACAG 120
AAGCTGATGG GCCAGATACA TCAGCTCAGA TCCGAATTAC AGGATATGGA GGCACAGCAA 180
10 GTTAATGAAG CAGAATCAGC AAGAGAACAG TTACAGGWTC TGCATGACCA AATAGCTGGG 240
CAGAAAGCAT CCAAACAAGA ACTAGAGACA GAACTGGAGC GACTGAAGCA GGAGTTCCAC 300
TATATAGAAG AAGATCTTTA TCGAACAAG AACACATTGC AAAGCAGAAT TAAAGATCGA 360
15 GACGAAGAAA TTCAAAACT CAGGAATCAG CITACCAATA AAACTTTAAG CAATAGCAGT 420
CAGTCTGAGT TAGAAAATCG ACTCCATCAG CTAACAGAGA CTCTCATCCA GAAACAGACC 480
20 ATGCTGGAGA GTCTCAGCAC AGAAAAGAAC TCCCTGGTCT TTCAACTGGA GCGCCTCGAA 540
CAGCAGATGA ACTCCGCTC TGAAGTAGT AGTAATGGGT CTTGATTAA TATGTCTGGA 600
ATTGACAATG GTGAAGGCAC TCGTCTGCGA AATGTTCTG TTTCTTTTAA TGACACAGAA 660
25 ACTAATCTGG CAGGAATGTA CGGAAAAGTT CGCAAAGCTG CTAGTTCAAT TGATCAGTTT 720
AGTATTCGCC TGGGAATTTT TCTCCGAAGA TACCCCATAG CGCGAGTTT TGTAATTATA 780
30 TATATGGCTT TGCTTACCT CTGGGTCATG ATTGTTCTGT TGAATTACAC ACCAGAAATG 840
CACCACGACC AACCATATGG CAAATGAACC AAGCCCAGTT GTTGCACTGA TTGGTTGTCT 900
TTTCTAGAC TTGGGATCTG CAAGAAGGCC AATGCGCTAA AATTCTGAG AACAGTGCAC 960
35 AAGATTATTT TATCACTACA AGCTTTTAAC TTTTAAAGTT ATTGTACAAG TATTCTACCT 1020
AAATCTTCCA ATTTCTTTA AATGGTAAGA GTTCTTAAAA CAGACAATAA TTTAACAAGC 1080
40 TCAGCTCTGC TTTATCTGAG TTTAGTGGTC CTAATATATA TGTAGAGAAA GATGGTGGGG 1140
TTGTTACCT CTGTACAGAC CATCTGTATG TTAGTGACA TTGATTATGG GTTATAATCA 1200
GGGAACTAA TTGTATTTAG TGACAAAAAT AAAAAGTTT TTTTATATA TTCAGTCTGC 1260
45 TTTTGGATTT TCATATATTT AACTTTGCAA AAAGATTAC TTTGTACATG TTACAGGCTT 1320
GATTGGTGTA AATCTTTTAA TAAATACATA AATAAAAGNA AAATATGCAT TTTCTTTTC 1380
50 TAAAAAAAAA AAAAAAAAAA CTCGA 1405

55 (2) INFORMATION FOR SEQ ID NO: 184:

(i) SEQUENCE CHARACTERISTICS:

- 60 (A) LENGTH: 1596 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 184:

5	GTCATGCACT GCGCCGAGA ACTGTGCTCT TTGAGGCCGA CGCTAGGGGC CCGGAAGGGA	60
	AACTGCGAGG CGAAGGTGAC CGGGGACCGA GCATTTTCAGA TCTGCTCGGT AGACCTGGTG	120
	CACCACCACC ATGTTGGCTG CAAGGCTGGT GTGTCTCCGG AACTACCTT CTAGGGTTTT	180
10	CCACCAGCT TTCACCAAG CCTCCCTGT TGTGAAGAAT TCCATCACGA AGAATCAATG	240
	GCTGTTAACA CCTAGCAGG AATATGCCAC CAAAACAAGA ATTGGGATCC GCGTGGGAG	300
15	AACTGGCCAA GAACTCAAAG AGGCAGCATT GGAACCATCG ATGGAAAAA TATTTAAAT	360
	TGATCAGATG GGAAGATGGT TTGTTGCTGG AGGGGCTGCT GTTGGTCTTG GAGCATTTG	420
	CTACTATGGC TTGGGACTGT CTAATGAGAT TGGAGCTATT GAAAAGGCTG TAATTTGGCC	480
20	TCAGTATGTC AAGGATAGAA TTCATTCCAC CTATATGTAC TTAGCAGGGA GTATTGGTTT	540
	AACAGCTTTG TCTGCCATAG CAATCAGCAG AACGCCGTGT CTCATGAACT TCATGATGAG	600
25	AGGCTCTTGG GTGACAATTG GTGTGACCTT TGCAGCCATG GTTGGAGCTG GAATGCTGGT	660
	ACGATCAATA CCATATGACC AGAGCCCAGG CCCAAAGCAT CTGCTTGGT TGCTACATTC	720
	TGGTGTGATG GGTGCAGTGG TGGCTCCTCT GACAATATTA GGGGTCCTC TTCTCATCAG	780
30	AGCTGCATGG TACACAGCTG GCATTGTGGG AGGCCTCTCC ACTGTGCCA TGTGTGCGCC	840
	CAGTGAAG TTTCTGAACA TGGGTGCACC CCTGGGAGTG GGCTGGGTC TCGTCTTTGT	900
35	GTCTCATTG GGATCTATGT TTCTTCCACC TACCACCGTG GCTGGTGCCA CTCTTTACTC	960
	AGTGGAATG TACGGTGGAT TAGTTCTTTT CAGCATGTTT CTCTGTATG ATACCCAGAA	1020
	AGTAATCAAG CGTGCAGAAG TATCACC AAT GTATGGAGTT CAAAAATATG ATCCCAITAA	1080
40	CTCGATGCTG AGTATCTACA TGGATACATT AAATATATTT ATGCGAGTTG CAACTATGCT	1140
	GGCAACTGGA GGCAACAGAA AGAAATGAAG TGAATCAGCT TCTGGCTTCT CTGCTACATC	1200
45	AAATATCTTG TTTAATGGG CAGATATGCA TTAAATAGTT TGTACAAGCA GCTTTCTGTTG	1260
	AAGTTTAGAA GATAAGAAAC ATGTCATCAT ATTTAAATGT TCCGGTAATG TGATGCCTCA	1320
	GGTCTGCCTT TTTTCTGGA GAATAAATGC AGTAATCCTC TCCCAAATAA GCACACACAT	1380
50	TTTCAATTCT CATGTTTGAG TGATTTTAAA ATGTTTGGT GAATGTGAAA ACTAAAGTTT	1440
	GTGTCAATG AATGTAAGTC TTTTCTTAC TTTAAATTT AGTAGGTTC CTGAGTAACT	1500
55	AAAATTTAGC AAACCTGTGT TTGCATATTT TTTKGGAGTG CAGMMTAWTG TAATTARAGC	1560
	ATTCCAGTAA NAGTGTMTT AAAGTTGNTC TATATN	1596

(2) INFORMATION FOR SEQ ID NO: 185:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 2293 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 185:

10 GCGCAGAGCC CGYACGAGCA GGACGACGAC GACAAGGGCG ACTCCAAGGA AACGCGGCTG 60
ACCCCTGATGG AGGAAGTGCT CCTGCTGGGC CTCAAGGACC GCGARGGTTA CACATCATTT 120
15 TGAATGACT GTATATCATC TGGATTACGT GGCTGTATGT TAATTGAATT AGCATTGAGA 180
GGAAGGTAC AACTAGAGGC TTGTGGAATG AGACGTAAAA GTCTATTAAAC AAGAAAGGTA 240
20 ATCTGTAAGT CAGATGCTCC AACAGGGGAT GTTCTTCTTG ATGAAGCTCT GAAGCATGTT 300
AAGGAACTC AGCCTCCAGA AACGGTCCAG AACTGGATTG AATTACTTAG TGGTGAGACA 360
TGGAATCCAT TAAAAATGCA TTATCAGTTA AGAAATGTAC GGGAACGATT AGCTAAAAAC 420
25 CTGGTGAAA AGGGTGTATT GACAACAGAG AAACAGAACT TCCTACTTTT TGACATGACA 480
ACACATCCCC TCACCAATAA CAACATTAAG CAGCGCCTCA TCAAGAAAGT ACAGGAAGCC 540
30 GTTCTTGACA AATGGGTGAA TGACCCTCAC CGCATGGACA GCGCTTGCT GGCCTCATT 600
TACCTGGCTC ATGCTCGGA CGTCTGGAG AATGCTTTTG CTCTCTTCT GGACGAGCAG 660
TATGATTTGG CTACCAAGAG AGTGGGCGAG CTTCTCGACT TAGACCCTGA AGTGAATGT 720
35 CTGAAGGCCA ACACCAATGA GGTCTGTGG GCGGTGGTGG CGGCGTTCAC CAAGTAACTC 780
TGCTCGGGGT GAACCATCTC CTTTCTCTC AAGTAAACCA GTAGTTTTTC TTCTGTTGAC 840
40 TTCTGGTTTT CTGTAATTTG TACTTTCCCA CACTATAATT GGCTTCTGTT TTACAAAATG 900
GTGGGTGGCT TTTTCTTTTT TGTACGTGA CAGGATCTG CTGGTACGAG AGGCCTTCCT 960
CTTTCTGTTT TAAAAAAG TTTTACTGCC ATATTGGCAT TCCATTCCCT GTTGCCATCC 1020
45 TCACGTGTTAC CTGTTTTGGG TTTCTGGTCT ACTTTGACTT TCAAAGTACC TCCAGCCTCC 1080
TCATACGCAC AGCTTTTGA TGACCTCAGC TTGAGTTTCT CCATATGTGC ATGTACATCT 1140
50 AGCATCTGC CTACAGTTCA GACAGAAGTC AAAAAAGGC CTTCAACTCA CCAAAGGTAA 1200
ATATCTGTAT CTATTAGGAC ATTTTTTACA TAGACTTCAG TTGAGATGTA TACTTAGCAA 1260
AATTATTTTT AAATTGAAAC AGCACAGTAA ATACTTAATA TAAATGTCC CTTGGATTTT 1320
55 GCTTCCCATG TAAATCTATT GTATTATTAC ACTTGTATA ATTTTAACTA TAAAGGTCCA 1380
ATTGTTTCAC AGAGCCAGTT TGGGATGGC TGCATTCCAT TTATGCTGTA TATAGTTTGA 1440
60 ATTATATATA AATTACCCCT TCTTCTGGCC ACCCTGCTC CCATCTTAGT ATTTTGCAAG 1500

ATCTAATCAG TTGTACACCT GGTGCCCTC GCTTGCTTCA ATCATGGTTA TTTGATGGCA 1560
AAATCGACCT CTTGTCGCTG AAGGAGAGAG AAAAGATGTG TGTCTGATTG GTCCTGGGAT 1620
5 TTTTGTAGCT GTGCCATTTA TGGTACTCTT TGCCTATGCA TCCCTTTTTT AGATTTTTTT 1680
TAAATTTTAT CTTACTGTTT TTATAATTTT TATTGGGAAG AGGCTTGTTA CCAGTACCAA 1740
10 TCTTGAGTTT CTTTTTCTGT CCACAAGTAA ATTAATATCT GCTCTGAAAT GTCATTTATC 1800
TACTCACACA TTCTTGGGGA AAAAAATCAA ATGTCAGTCC TAGCAGATGT TGCATGTAAA 1860
TTGGTAGCAA GTAATGATTA CAACCCAGAG GATTAAGAAT TTTGTAACAG AAAGCTCTAT 1920
15 GTTTTAATTT TTTATATACA ATTAGGATAA TTAGCATTGT CAGACTATAA ACCTTTGCTT 1980
TTTAAAGTTT ATTTTACTA TTTCTTTATC ACTTTATTGT ATCATCACCA TTGGTTTCAT 2040
20 AATGTAAATA CTATATGTTG AACAAATTAA ATGTCAAAT TTTTATTAC CATAGTCCAT 2100
GTTAATAGTG GGGCTTCAG GTGTTAGAG ATTTTMTTG TTGTTGTAA CATTCATTGC 2160
AAAAGTACTA GATGGTGAT AACTCTAGAG TTGAATTTTA AGGGATTCCC TAATATGTAT 2220
25 ACTATCTTTT TATCTGAAGT AATAAATAA CAATGATCTT GAAAGTCCY RAAAMAAAAA 2280
AAAAAAAAA AAA 2293

30

(2) INFORMATION FOR SEQ ID NO: 186:

35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1212 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
40 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 186:

GGCAGGAGGC GAGCCGGCGC ACCGTACGCT GGGACGTGTG GTTTCAGCTC GTGCGCCTCC 60
45 CCGTGGGTTT GCGACGTTTA GCGACTATTG CGCCTGCGCC ACGCCGGCTG CGAGACTGGG 120
GCCGTGGCTG CTGGTCCGG GTGATGCTAG GCGGCTCCCT GGGCTCCAGG CTGTTGCGGG 180
GTGTAGGTGG GAGTCACGGA CGGTTCGGGG CCCGAGGTGT CCGCGAAGGT GGCGCACATG 240
50 GGCGGCAGGG GAGAGCATGG CTCAGCGGAT GGTCTGGGTG GACCTGGAGA TGACAGGATT 300
GGACATGAG AAGGACCAGA TTATTGAGAT GGCCTGTCTG ATAAGTACT CTGATCTCAA 360
55 CATTTTGGCT GAAGGTCCTA ACCTGATTAT AAAACAACCA GATGAGTTGC TGGACAGCAT 420
GTCAGATGG TGTAAGGAGC ATCAGGGGAA GTCTGGCCTT ACCAAGGCAG TGAAGGAGAG 480
TACAATTACA TTGCAGCAGG CAGAGTATGA ATTTCTGTCC TTTGTACGAC AGCAGACTCC 540
60

TCCAGGGCTC TGTCCACTTG CAGGAAATTC AGTTCATGAA GATAAGAAGT TTCTTGACAA 600
ATACATGCCC CAGTTCATGA AACATCTTCA TTATAGAATA ATTGATGTGA GCACTGTATA 660
5 AGAACTGTGC AGACGCTGGT ATCCAGAAGA ATATGAATTT GCACCAAAGA AGGCTGCTTC 720
TCATAGGGCA CTTGATGACA TTAGTGAAAG CATCAAAGAG CTTCAAGTTT ACCGAAATAA 780
CATCTTCAAG AAAAAAATAG ATGAAAAGAA GAGGAAAATT ATAGAAAATG GGGAAAATGA 840
10 GAAGACCGTG AGTTGATGCC AGTTATCATG CTGCCACTAC ATCGTTATCT GGAGGCAACT 900
TCTGGTGGTT TTTTCTCTC ACGCTGATGG CTTGGCAGAG CACCTTCGGT TAACTTGCAT 960
15 CTCCAGATTG ATTACTCAAG CAGACAGCAC ACGAAATACT ATTTTCTCC TAATATGCTG 1020
TTTCCATTAT GACACAGCAG CTCCTTTGTA AGTACCAGGT CATGTCCATC CCTTGGTACA 1080
TATATGCATT TGCTTTTAAA CCATTTCTTT TGTTTAAATA AATAAATAAG TAAATAAAGC 1140
20 TAGTTCTATT GAAATGCAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1200
AAAAAAAAA AN 1212

25

(2) INFORMATION FOR SEQ ID NO: 187:

- 30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1605 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
35 (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 187:

GCTTCCGAA GTTGCTTTTG TCCAACATC CGGGCTTCTC CTTTGTGTG TCCGGCCGAT 60
40 CCCACCTCTC CTCGACCTG GACGTCTACC TTCCGGAGGC CCACATCTTG CCCACTCCGC 120
GCGCGGGCT AGCGCGGGT TCAGCGACGG GAGCCCTCAA GGGACATGGC AACTACAGCG 180
GCGCCGGCG GCGCGCCCG AAATGGAGCT GGCCCGAAT GGGGAGGGT CGAAGAAAAC 240
45 ATCCAGGGCG GAGGCTCAGC TGTGATTGAC ATGGAGAACA TGGATGATAC CTCAGGCTCT 300
AGCTTCGAGG ATATGGGTGA GCTGCATCAG CGCCTGCCG AGGAAGAAGT AGACGCTGAT 360
50 GCAGCTGATG CAGCTGCTGC TGAAGAGGAG GATGGAGAGT TCCTGGGCAT GAAGGCTTT 420
AAGGGACAGC TGAGCCGCA GGTGGCAGAT CAGATGTGGC AGGCTGGGAA AAGACAAGCC 480
TCCAGGGCCT TCAGCTTGTG CGCCAACATC GACATCCTCA GACCCTACTT TGATGTGGAG 540
55 CCTGCTCAGG TGCGAACAGG GCTCCTGGAG TCCATGATCC CTATCAAGAT GGTCAACTTC 600
CCCCAGAAA TTGCAGGTGA ACTCTATGGA CCTCTCATGC TGGTCTTCAC TCTGGTTGCT 660
60 ATCTACTCC ATGGGATGAA GACGTCTGAC ACTATTATCC GGGAGGGCAC CCTGATGGGC 720

ACAGCCATTG GCACCTGCTT CGGCTACTGG CTGGGAGTCT CATCCTTCAT TTACTTCCTT 780
5 GCCTACCTGT GCAACGCCCA GATCACCATG CTGCAGATGT TGGCACTGCT GGGCTATGGC 840
CTCTTTGGGC ATTGCAITGT CCTGTTTCATC ACCTATAATA TCCACCTCCA CGCCCTCTTC 900
TACCTCTTCT GGCTGTTGGT GGGTGGACTG TCCACACTGC GCATGGTAGC AGTGTGTTGGT 960
10 TCTCGGACCG TGGGCCCCAC ACAGCGGCTG CTCCTCTGTG GCACCTGGC TGCCCTACAC 1020
ATGCTCTTCC TGCTCTATCT GCATTTTGCC TACCACAAAG TGGTAGAGGG GATCCTGGAC 1080
ACACTGGAGG GCCCCAACAT CCCGCCCATC CAGAGGGTCC CCAGAGACAT CCTTGCCATG 1140
15 CTCCTGCTG CTCGGCTTCC CACCACCGTC CTCAACGCCA CAGCCAAAGC TGTGCGGTG 1200
ACCTTGCACT CACACTGACC CCACCTGAAA TTCTTGCCCA GTCTCTTTTC CCGCAGCTGC 1260
20 AGAGAGGAGG AAGACTATTA AAGGACAGTC CTGATGACAT GTTTCGTAGA TGGGGTTTGC 1320
AGCTGCCACT GAGCTGTAGC TGCCTAAGTA CCTCCTTGAT GCNTGTGGC ACTTCTGAAA 1380
GGCACAAGGC CAAGAACTCC TGGCCAGGAC TGCAAGGCTC TGCAGCCAAT GCAGAAAATG 1440
25 GGTGAGCTCC TTTGAGAACC CCTCCCCACC TACCCCTTCC TTCTCTTTTA TCTCTCCAC 1500
ATTGTCTTGC TAAATATAGA CTGGTAATT AAAATGTTGA TTGAAGTCTG GAAAAAATA 1560
30 AAAAAAATA AAAAAAATA AAAAAAATA AAAAAAAT TCGAG 1605

35 (2) INFORMATION FOR SEQ ID NO: 188:

(i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 1516 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 188:

45 ATTCGGCATG AGGGGGTCAC GTGGTGGCTG GGCCGGGGAA ATGGCGGCTT CAGGAGAGAG 60
CGGGACTTCA GCGGGCGGAG GCAGCACCAG GGAAGCATTT ATGACCTTCT ACAGTGAGGT 120
50 GAAACAAATA GAGAAGAGAG ACTCGGTCTT AACTTCGAAA AATCAGATG AAAGACTGAC 180
CCGTCCCTGGT TCCTCTTACT TCAATTTGAA CCCATTTGAG GTTCTTCAGA TAGATCCTGA 240
AGTTACAGAT GAAGAAATAA AAAAGAGGTT TCGGCAGTTA TCCATCTTGG TGCATCCTGA 300
55 CAAAAATCAA GATGATGCTG ACAGAGCACA AAAGGCTTTT GAAGCTGTGG ACAAAGCTTA 360
CAAGTTGCTA CTGGATCAGG AGCAAAAGAA GAGGGCCCTG GATGTAATTC AGGCAGGAAA 420
60 AGAATACGTG GAACACACTG TGAAAGAGCG AAAAAACAA TTAAAGAAGG AAGGAAAACC 480

TACAATTGTA GAGGAGGATG ATCCTGAGCT GTTCAAACAA GCTGTATATA AACAGACAAT 540
GAAACTCTTT GCAGAGCTGG AAATTAAAAG GAAAGAGAGA GAAGCCAAAG AGATGCATGA 600
5 AAGGAAACGA CAAAGGGAAG AAGAGATTGA AGCTCAAGAA AAAGCCAAAC GGGAAAGAGA 660
GTGGCAGAAA AACTTTGAGG AAAGTCGAGA TGGTCGTGTG GACAGCTGGC GAAACTTCCA 720
AGCCAATACG AAGGGGAAGA AAGAGAAGAA AAATCGGACC TTCCTGAGAC CACCGAAAGT 780
10 AAAAATGAG CAACGTGAGT GACCGCCCAA GGTCACAGGC ACAGAACCTT TCCCCTGCTA 840
TCTCCCTTCC TGCTTCGAAG GACTCATTCT TTCCTCCAC TTCCACCCCA ACATAGAGTA 900
15 GTATTTGCTT TTTAGTCCAT TTTGTTTTCA ATACGATTTA ATATCGATCA GAGTAATTCT 960
TTTGTACATT GAAATGAGGG GCTTGGTTTA AAAAAAGACC TTCCCTCTC CTGCCCCTA 1020
GAACAACCAG TATTAGAAGG TGCCACCATT GGTGCTGCCT TCTCTTCCA CAGCCTGTAA 1080
20 CTCAGTGTTC TGTACTTCAC TGAATTGTGA TGGTTAGAAA CTTCGTGGAT AGTTTGTGGA 1140
AATCATCCAA TTAAACATAC TGCTTAAAC AGTGTGCTG TGACTTCAGA GACAAGCCTG 1200
25 GAAGGGGCAC CTTAGGAAGC CCCTTCGCTT CAGTTGCTCG CTCTGGGTG TGCTCCCTTC 1260
GAAGGCCAG ATAAGACAGG GAACACTTGT GAGCACACAG AGCAGCATCT GATGCCCTGT 1320
GGTGTTTGGC ATGTGCCCC TGCTACTGA CCAATCAGTG TGGCATGAGG CCCACGCCAC 1380
30 CCAAACCTTT CACTTTCCAA AGAGCTAGCC GTCCCTCACC CAGTACCATG TCCTAGCCTG 1440
TCTGCATTTG TTAGTGGTAA TATTCTTTAT GTATAATAAA TTTTATACC CAAAAAAAAA 1500
35 AAAAAAAAAA ACTCGA 1516

40 (2) INFORMATION FOR SEQ ID NO: 189:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 681 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 189:

50 GCTCCCATGT TGCTGGCTGT CCGTACATCA CCCTGTCCCC TGCAGGAGGG GGCTACAGGC 60
CATCTCCCTC CTGTAGGCCT CTGACTCCCC TCCACTTTTG GGCCTCAGC TTATCTGGG 120
CAGGGACCA TTGCAGCATC CTCCCTCCT CNGGACTCAA GGTGCTGAGG TATAAGCCCT 180
55 GGGCCCCAGA TCCCTGRTKA CACCTTCCTG GAGAAGACTC TCAAAGTGA CTGTATATTT 240
GAGTTCACCA GCAATAACTC CCCCACTCG AAGCAGGTCC AAACCMAGG ATCCCAGGGT 300
60 CCTTGGGCTC TGTGGCACTG TCTTCCCAAG ATCTTCCTG TTGCACAATG GGAAACCTAA 360

5 GAGGAAAAAG ACAGGGGCCT GCTTGCCCAG CCATGCGAGG GATTCCATGC CCACCTGCCC 420
TCTGYCTGCC TCGCTGGAAT GTGGGCCCCCT GCTCCCCGTC AGGTGTGTGCT GTCTCTGACC 480
TATGTTTACA TCCCCGAGGG GTTCTGCCT CCTCCCCACC CAGGTCAGGG TGTGGTCCAG 540
CAGCTTGCTG TGGGGTGCTG ACATGTGTCA CCACTGCCCC CCTTGCCCCC GGGGGGGTCA 600
10 TGGTCTCCTC CTGGATGCTG CTCCTTGAAT YTTTTTYYT GAWAAACCYT TTAMAATTAA 660
AAAAAAAAA AAAAAACTCG A 681

15

(2) INFORMATION FOR SEQ ID NO: 190:

20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1014 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 190:

GCCTCAAGCC ACGCATATGA TAATTTCTG GAACATCAA ATTCAGTGTT TCTACAGCCA 60
GTTAGTCTAC AAACCATTCG AGCAGCACCA TCAAACCAGA GTCTGCCACT TTTTGTGTCATC 120
30 CCTGGATGAT TGCTGGGCAA AGGTGGCCTT TTAGAGCTCT TAAAAGCCCA CAAAAAGGCT 180
ATTCGTAGAG CCACAGTCAA CACATTTGGT TATATTGCAA AGGCCATTGG CCTCATGATG 240
35 TATTGGCTAC ACTTCTGAAC AACCTCAAAG TTCAAGAAAG GCAGAACAGA GTTTGTACCA 300
CTGTAGCAAT AGCTATTGTT GCAGAAACAT GTTCACCCCT TACAGTACTC CCTGCCTTAA 360
40 TGAATGAATA CAGAGTTCCT GAACTGAATG TTCAAATGG AGTGTTAAAA TCGCTTTCCT 420
TCTTGTTTGA ATATATTGGT GAAATGGGAA AAGACTACAT TTATGCCGTA ACACCGTTAC 480
TTGAAGATGC TTAAATGGAT AGAGACCTTG TACACAGACA GACGGCTAGT GCAGTGGTAC 540
45 AGCACATGTC ACTTGGGGTT TATGGATTTG GTTGTAAGA TTCGCTGAAT CACTGTGTGA 600
ACTATGTATG GCCCAATGTR TTTGAGACAT CTCCTCATGT AATTCAGGCA GTTATGGGAG 660
CCCTAGAGGG CCTGAGAGTT GCTATTGGAC CATGTAGAAT GTTGCAATAT TGTTTACAGG 720
50 GTCTGTTTCA CCCAGCCCGG AAAGTCAGAG ATGTATATTG GAAAATTAC AACTCCATCT 780
ACATGGGTC CCAGGACGCT CTCATAGCAC ATTACCCAAG AATCTACCAA CGATGATAAG 840
55 RACACCTATA TTCGTTATGA ACTTGACTAT ATCTTATAAT TTTATTGTTW ATTTKGTKT 900
TAATGCACAS TACTTCACAC CTTAACTTG CTTTGATTG GTGATGTAAA CTTTTAAACA 960
TTGCAGATCA GTGTAGGACT GGTCCATAGG GGAAGAGCTA GGAANTCCAT AGGC 1014
60

(2) INFORMATION FOR SEQ ID NO: 191:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2779 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 191:

5	TCGCAGCAGG GTGTGTCCAG ATGGTCAGTC TCTGGTGGCT AGCCTGTCCT GACAGGGGAG	60
15	AGTTAAGCTC CCGTCTCCA CCGTGCCGCG TGGCCAGGTG GGCTGAGGGT GACCGAGAGA	120
	CCAGAACCTG CTTGCTGGAG CTTAGTGCTC AGAGCTGGGG AGGAGGTTTC CGCCGCTCCT	180
20	CTGCTGTGCG CGCCGGCAGC CCTTCCCGGC TTCACTTCCT CCCGAGGCC CTGCTACTGA	240
	GAAGCTCCGG GATCCCAGCA GCCGCCACGC CCTGGCCTCA GCCTGCGGGG CTCCAGTCAG	300
25	GCCAACACCG ACGCGCANTG GGAGGAAGAC AGGACCCTTG ACATCTCCAT CTGCACAGAG	360
	GTCTTGCGTG GACCGAGCAG CCTCTCTCTC CTAGGATGAC CTCACCTCC AGCTCTCCAG	420
	TTTTCAGGTT GGAGACATTA GATGGAGGCC AAGAAGATGG CTCTGAGGCG GACAGAGGAA	480
30	AGCTGGATTT TGGGAGCGGG CTGCCTCCCA TGGAGTCACA GTTCCAGGGC GAGGACCGGA	540
	AATTCGCCCC TCAGATAAGA GTCAACCTCA ACTACCGAAA GGAACAGGT GCCAGTCAGC	600
35	CGGATCCAAA CCGATTGAC CGAGATCGGC TCTTCAATGC GGTCTCCCGG GGTGTCCCCG	660
	AGGATCTGGC TGGACTTCCA GAGTACCTGA GCAAGACCAG CAAGTACCTC ACCGACTCGG	720
	AATACACAGA GGGCTCCACA GGTAAAGCGT GCCTGATGAA GGCTGTGCTG AACCTTAAGG	780
40	ACGGGGTCAA TGCCTGCATT CTGCCACTGC TGCAGATCGA CCGGGACTCT GGCAATCCTC	840
	AGCCCCTGGT AAATGCCCAG TGCACAGATG ACTATTACCG AGGCCACAGC GCTCTGCACA	900
45	TGCCCATGTA GAAGAGGAGW CTGCAGTGTG TGAAGCTCCT GGTGGAGAAT GGGGCCAATG	960
	TGCATGCCCC GGTCTGCGGC GCTTCTTCCA GAAGGGCCAA GGGACTTGCT TTTATTTCCG	1020
	TGAGCTACCC CTCTYTTTGG CCGCTTGAC CAAGCAGTGG GATGTGGTAA GCTACCTCCT	1080
50	GGAGAACCCA CACCAGCCCG CCAGCCTGCA GGCCTGACT CCCAGGGCAA CACAGTCTTG	1140
	CATGCCCTAG TGATGATCTC GGACAACTCA GCTGAGAACA TTGCACTGGT GACCAGCATG	1200
55	TATGATGGGC TCCTCCAAGC TGGGGCCCGC CTCTGCCCTA CCGTGCAGCT TGAGGACATC	1260
	CGCAACCTGC AGGATCTCAC GCCTCTGAAG CTGGCCGCCA AGGAGGGCAA GATCGAGATT	1320
	TTCAGGCACA TCCTGCAGCG GGAGTTTTC A GACTGAGCC ACCTTTCCCG AAAGTTCACC	1380
60	GAGTGGTGCT ATGGGCCTGT CCGGGTGTCG CTGTATGACC TGGCTTCTGT GGACAGCTGT	1440

GAGGAGAACT CAGTGCTGGA GATCATTGCC TTTCATTGCA AGAGCCCGCA CCGACACCGA 1500
ATGGTCGPTT TGGAGCCCCCT GAACAACTG CTGCAGGCGA AATGGGATCT GCTCATCCCC 1560
5 AAGTCTTCTT TAAACTTCCT GTGTAATCTG ATCTACATGT TCATCTTCAC CGCTGTTGCC 1620
TACCATCAGC CTACCCTGAA GAAGCAGGCC GCCCCTCACC TGAAAGCGGA GGTGGAAC 1680
10 TCCATGCTGC TGACGGGCCA CATCCTTATC CTGCTAGGGG GGATCTACCT CCTCGTGGGC 1740
CAGCTGTGGT ACTTCTGGCG GCGCCACGTG TTCATCTGGA TCTCGTTCAT AGACAGCTAC 1800
TTTGAAATCC TCTTCTGTT CCARGCCCTG CTCACAGTGG TGTCCARGT GCTGTGTTTC 1860
15 CTGGSCATCG AGTGGTACCT GCCCCTGCTT GTGTCTGCGC TGGTGCTGGG CTGGCTGAAC 1920
CTGCTTTACT ATACACGTGG CTTCAGCAC ACAGGCATCT ACAGTGTCTAT GATCCAGAAG 1980
20 CCCTGGTGAG CCTGAGCCAG GANNTTGGCG CCCCAGAGCT CCTACAGGCC CCAATGCCAC 2040
AGAGTCAGTG CAGCCCATGG AGGGACAGGA KGACGAKGGC AACGGGGCCC AGTACAGGGG 2100
TATCCTGGAA GCCTCCTTGG AGCTCTTCAA ATTACCATC GGCATGGGCG AGCTGGCCTT 2160
25 CCAGGARCAG CTGCACCTCC GCGGCATGGT GCTGCTGCTG CTGCTGGSCT ACGTGTCTGT 2220
CACCTACATC CTGCTGCTCA ACATGCTCAT CGCCCTCATG AGCGAGACCG TCAACAGTGT 2280
30 CGCCACTGAC AGCTGGAGCA TCTGGAAGCT GCAGAAAGCC ATCTCTGTCC TGGAGATGGA 2340
GAATGGCTAT TGGTGGTGCA GGAAGAAGCA GCGGGCAGGT GTGATGCTGA CCGTTGGCAC 2400
TAAGCCAGAT GGCAGCCCSG ATGAGCGCTG GTGCTTCAGG GTGGAGGAGG TGAACGGGC 2460
35 TTCATGGGAG CAGACGCTGC CTACGCTGTG TGAGGACCCG TCAGGGGCAG GTGTCCCTCG 2520
AACTCTCGAG AACCTGTGCC TGGCTTCCCC TCCCAAGGAG GATGAGGATG GTGCCTCTGA 2580
40 GGAAACTAT GTGCCGTCC AGCTCCTCCA GTCCAACTGA TGGCCAGAT GCAGCAGGAG 2640
GCCAGAGGAC AGAGCAGAGG ATCTTTCCAA CCACATCTGC TGGCTCTGGG GTCCCACTGA 2700
ATTCTGGTGG CAAATATATA TTTTCACTAA CTCAAAAAAA AAAAAAAAAA AAAAAAAAAA 2760
45 AAAAAAAAAA AAAAAAGGC 2779

50

(2) INFORMATION FOR SEQ ID NO: 192:

(i) SEQUENCE CHARACTERISTICS:

55

- (A) LENGTH: 1923 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 192:

	ACCCGCTCCG CTCGCTCCG CTCGGCCCCG CGCGCCCGT CAACATGATC CGCTGCGGCC	60
	TGGCCTGCGA GCGCTGCCG TGGATCCTGC CCCTGCTCCT ACTCAGCGCC ATCGCCTTCG	120
5	ACATCATCGC GCTGGCCGGC CGCGGCTGGT TGCAGTCTAG CGACCACGGC CAGACGTCCT	180
	CGCTGTGGTG GAAATGCTCC CAAGAGGGCG GCGGCAGCGG GTCCTACGAG GAGGGCTGTC	240
10	AGAGCCTCAT GGAGTACGCG TGGGGTAGAG CAGCGGCTGC CATGCTCTTC TGTGGCTTCA	300
	TCATCCTGGT GATCTGTTTC ATCCTCTCCT TCTTCGCCCT CTGTGGACCC CAGATGCTTG	360
	TCTTCCTGAG AGTGATGGA GGTCTCCTTG CCTTGGCTGC TGTGTTCCAG ATCATCTCCC	420
15	TGTTAATTTA CCCCGTGAAG TACACCCAGA CCTTCACCCT TCATGCCAAC CSTGCTGTCA	480
	CTTACATCTA TAACTGGGCC TACGGCTTTG GGTGGGCAGC CACGATTATC CTGATYGGCT	540
20	GTGCCCTCTT CTCTGCTGC CTCCCCAACT ACGAAGATGA CCTTCTGGGC AATGCCAAGC	600
	CCAGGTACTT CTACACATCT GCCTAACTTG GGAATGAATG TGGGAGAAAA TCGCTGCTGC	660
	TGAGATGGAC TCCAGAAGAA GAAACTGTTT CTCCAGGCGA CTTTGAACCC ATTTTGTGGC	720
25	AGTGTTCATA TTATTAACT AGTCAAAAAT GCTAAAATAA TTTGGGAGAA AATATTTTTC	780
	AAGTAGTGT ATAGTTTCAT GTTTATCTTT TATTATGTTT TGTGAAGTTG TGTCTTTTCA	840
30	CTAATTACCT ATACTATGCC AATATTTCCT TATACTATC CATAACATTT ATACTACATT	900
	TGTAAGAGAA TATGCACGTG AAACCTAACA CTTTATAAGG TAAAAATGAG GTTTCCAAGA	960
	TTTAATAATC TGATCAAGTT CTGTATTATT CCAATAGAA TGGACTCGGT CTGTAAAGG	1020
35	CTAAGGAGAA GAGGAAGATA AGTTTAAAG TTGTTAATGA CCAACATTC TAAAAGAAAT	1080
	GCAAAAAAA AGTTTATTTT CAAGCCTTCG AACTATTTAA GGAAAGCAAA ATCATTTCCT	1140
40	AAATGCATAT CATTTGTGAG AATTTCTCAT TAATATCCTG AATCATTCAT TTCAGCTAAG	1200
	GCTTCATGTT GACTCGATAT GTCATCTAGG AAAGTACTAT TTCATGGTCC AAACCTGTTG	1260
	CCATAGTTGG TAAGGCTTTC CTTTAAGTGT GAAATATTTA GATGAAATTT TCTCTTTTAA	1320
45	AGTTCITTAT AGGGTAGGG TGTGGGAAA TGCTATATTA ATAAATCTGT AGTGTTTTGT	1380
	GTTTATATGT TCAGAACCAG AGTAGACTGG ATTGAAAGAT GGACTGGGTC TAATTTATCA	1440
50	TGACTGATAG ATCTGGTTAA GTGTGTAGT AAAGCATTAG GAGGGTCATT CTTGTCACAA	1500
	AAGTGCCACT AAAACAGCCT CAGGAGAATA AATGACTTGC TTTTCTAAAT CTCAGGTTTA	1560
	TCTGGGCTCT ATCATATAGA CAGGCTTCTG ATAGTTTGCA ACTGTAAGCA GAAACCTACA	1620
55	TATAGTTAAA ATCCTGGTCT TTCTTGGTAA ACAGATTTTA AATGTCTGAT ATAAAACATG	1680
	CCACAGGAGA ATTCGGGGAT TTGAGTTTCT CTGAATAGCA TATATATGAT GCATCGGATA	1740
60	GGTCATTATG ATTTTACC ATTTGACTT ACATAATGAA AACCAATTCA TTTTAAATAT	1800

	CAGATTATTA TTTTGTAAGT TGTGGAAAA GCTAATTGTA GTTTTCATTA TGAAGTTTTC	1860
	CCAATAAACC AGGTATTCTA AAAAAAAAAA AAAAAAACTN GAGGGGGGGC CCGGTACCCA	1920
5	ATT	1923
10	(2) INFORMATION FOR SEQ ID NO: 193:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 2346 base pairs	
	(B) TYPE: nucleic acid	
15	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 193:	
20	AGGCTCAGGG GGACACTCTC AAAATTACAC AGCTTTTAAC AGGTGGCAGA ATTGGGGTTC	60
	AGACCCAGAT CTGGGTTCAA GTCACATG GTGTGATTGC GGCATTCTTT CCCGCATCTG	120
	GGCCTTGCCA TCTCTCTCTC CGAGTGGACA TGGAGAGGAC GGGGGCCCAG CAGCTGGATG	180
25	GCTGCAGGG ATCAAGTCTT CTCTGGGCT GGGCACGTAN AAGAGCATGT GGCTGGTGA	240
	CGGCATGCTT GGCTCCTCAC CTGGCAGTCT GCCTGCCCTG CTAACCGGCT GTCTCTGTGT	300
30	CCCCTAGTGC CCTGGGCTAG CATGACCCGC CTGATGCGWT SCCGCACAGC CTCTGGTTCC	360
	AGCGTCATTC TCTGGATGGC ACCCGCAGCC GCTCCACAC CAGCGAGGGC ACCCGAAGCC	420
	GCTCCACAC CAGCGAGGGC ACCCGCAGCC GCTCGCACAC CAGCGAGGGG GCCCACCTGG	480
35	ACATCACCCC CAACTCGGGT GCTGCTGGGA ACAGNGCCCG GCCCAAGTCC ATGGAGGTCT	540
	CCTGCTAGGC GGCCTGCCCA GCTGCCGCC CCGGACTCTG ATCTCTGTAG TGGCCCCCTC	600
40	CTCCCCGGCC CCTTTTCGCC CCCTGCCTGC CATACTGCGC CTAACCTGGT ATTAATCCAA	660
	AGCTTATTTT GTAAGAGTGA GCTCTGGTGG AGACAAATGA GGTCTATTAC GTGGGTGCCC	720
	TCTCCAAAGG CGGGGTGGCG GTGGACCAA GGAAGGAAGC AAGCATCTCC GCATCGCATC	780
45	CTCTTCCATT AACCAGTGGC CGGTGCCCAC TCTCTCCCC TCCCTCAGAG ACACCAAAT	840
	GCCAAAAACA AGACCGGTAC AGCACACACT TCACAAAGCC AAGCCTAGGC CGCCCTGAGC	900
50	ATCCTGGTTC AAACGGGTGC CTGGTCAGAA GGCCAGCCGC CCACTTCCCG TTTCTCTTT	960
	AACTGAGGAG AAGCTGATCC AGTTTCCGGA AACAAAATCC TTTTCTCATT TGGGGAGGGG	1020
	GGTAATAGTG ACATGCAGGC ACCTCTTTTA AACAGGCAAA ACAGGAAGGG GGAAAAGGTG	1080
55	GGATTCATGT CGAGGCTAGA GGCATTTGGA ACAACAAATC TACGTAGTTA ACTTGAAGAA	1140
	ACCGATTTTT AAAGTTGGTG CATCTAGAAA GCTTTGAATG CAGAAGCAAA CAAGCTTGAT	1200
60	TTTCTAGCA TCCTCTTAAT GTGCAGCAAA AGCAGGCRAC AAAATCTCCT GGCTTTACAG	1260

ACAAAAATAT TTCAGCAAAC GTTGGGCATC ATGGTTTTTG AAGGCTTTAG TTCTGCTTTC 1320
TGCTCTCCT CCACAGCCCC AACCTCCAC CCCTGATACA TGAGCCAGTG ATTATCTTG 1380
5 TTCAGGGAGA AGATCAITTA GATTGTGTTT GCATTCTTA GAATGGAGG CAACATTCCA 1440
CAGCTGCCCT GGCTGTGATG AGTGTCTTG CAGGGGCCG AGTAGAGCA CTGGGTGGG 1500
10 GCGGAATTG GGGTTACTCG ATGTAAGGA TTCTTGTG TGTTGTGAG ATCCAGTGCA 1560
GTGTGATTT CTGTGGATCC CAGCTGGTT CCAGGAATTT TGTGTGATT GCTTAAATCC 1620
AGTTTCAAT CTTGACAGC TGGCTGGAA CGTGAATCA GTAGCTGAAC CTGTCTGACC 1680
15 CGGTACGTT CTGGATCCT CAGAACTCT TGCTCTGTC GGGTGGGGG TGGGAATCA 1740
CTGGGGAGC GGTGGCTGAG AAAATGTAAG GATTCTGAA TACATATTCC ATGGGACTTT 1800
20 CCTCCCTCT CTGCTTCCT CTTTCTCTGC TCCCTAACCT TTCGCCAAT GGGGCAGCAC 1860
CACTGACGTT TCTGGGCGC CAGTGGGCT GCCAGGTTCC TGTACTACTG CCTGTACTT 1920
TTCATTTGG CTCACCGTG ATTTCTCAT AGGAAGTTG GTCAGAGTGA ATTGAATATT 1980
25 GTAAGTCAGC CACTGGGACC CGAGATTTC TGGACCCCG CAGTTGGGAG GAGGAAGTAG 2040
TCCAGCCTC CAGGTGGCGT GAGAGGCAAT GACTCGTTAC CTGCCGCCA TCACCTTGA 2100
30 GGCCTTCCCT GGCCTTGAGT AGAAAAGTCG GGGATCGGG CAAGAGAGC TGAGTACGA 2160
TGGGAACTA TTGTGCACAA GTCTTCCAG AGGAGTTCT TAATGAGATA TTTGTATTTA 2220
TTCCAGACC AATAAATTG TAACTTTGCA AAAAAAAAA AAAAAAAAA AAAAAAAAA 2280
35 AAAAAAAAA AAAAAAACT CGAGGGGGC CCGTACCAA TTCGCCGTAT ATGATCGTAA 2340
ACAATC 2346
40

(2) INFORMATION FOR SEQ ID NO: 194:

- 45 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3054 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
50
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 194:

TATCTGAACC ACCCTTTATT CTACATATGA TAGGCAGCAC TGAAATATCC TAACCCCTA 60
55 AGCTCMAGGT GCCCTGTGGN ACGAGCAACT GGAATATAGC AGGGCTGGG TCTGTCTTCC 120
TGGTATAGG CTCACTCTT CCCCCAATC TTCTCTGGA GCTTTCAGC CAAGTGCTA 180
AAAGGAATAG GTAGGAGACC TCTCTATCT AATCCTTAA AGCATAATGT TGAACATTCA 240
60

	TTCAACAGCT GATGCCCTAT AACCCCTGCC TGGATTTCCTT CCTATTAGGC TATAAGAAGT	300
	AGCAAGATCT TTACATAATT CAGAGTGGTT TCATTGCCTT CCTACCCTCT CTAATGGCCC	360
5	CTCCATTTAT TTGACTAAAG CATCACACAG TGGCACTAGC ATTATACCAA GAGTATGAGA	420
	AATACAGTGC TTTATGGCTC TAACATTACT GCCTTCAGTA TCAAGGCTGC CTGGAGAAAG	480
	GATGGCAGCC TCAGGGCTTC CTTATGTCCT CCACCACAAG AGCTCCTTGA TGAAGGTCAT	540
10	CTTTTTCCTCC TATCCTGTTT TCCCCCTCCC CGCTCCTAAT GGTACGTGGG TACCCAGGCT	600
	GGTCTTGGG CTAGGTAGTG GGGACCAAGT TCATTACCTC CCTATCAGTT CTAGCATAGT	660
15	AAACTACGGT ACCAGTGTTA GTGGGAAGAG CTGGGTTTTT CTAGTATACC CACTGCATCC	720
	TACTCCTACC TGGTCAACCC GCTGCTTCCA GGTATGGGAC CTGCTAAGTG TGAATTACC	780
	TGATAAGGGA GAGGGAAATA CAAGGAGGGC CTCTGGTGTT CTGGCCTCA GCCAGCTGCC	840
20	CACAAGCCAT AAACCAATAA AACAAGAATA CTGAGTCAGT TTTTATCTG GGTCTCTTC	900
	ATTCCCCTG CACTTGGTGC TGCTTGGCT GACTGGGAAC ACCCCATAAC TACAGAGTCT	960
25	GACAGGAAGA CTGGAGACTG *TCCACTTCTA GCTCGGAAC TACTGTGTAA ATAAACTTTC	1020
	AGAACTGCTA CCATGAAGTG AAAATGCCAC ATTTTGCTTT ATAATTCTA CCCATGTTGG	1080
	GAAAACTGG CTTTTTCCCA GCCCTTCCA GGGCATAAAA CTCAACCCCT TCGATAGCAA	1140
30	GTCCCATCAG CCTATTATTT TTTTAAAGAA AACTTGCACT TGTTCCTTT TTTACAGTTA	1200
	CTTCCTTCCT GCCCAAAAT TATAAACTCT AAGTGTAAAA AAAAGTCTTA ACAACAGCTT	1260
35	CTTGCTTGTA AAAATATGTA TTATACATCT GTATTTTAA ATCTGCTCC TGAAAAATGA	1320
	CTGTCCCAT CTCCACTCAC TGCATTGGG GCCTTTCCA TTGGTCTGCA TGTCTTTTAT	1380
	CATTGCAGGC CAGTGACAG AGGGAGAAGG GAGAACAGGG GTCGCCAACA CTTGTGTTGC	1440
40	TTTCTGACTG ATCCTGAACA AGAAAGAGTA ACACTGAGGC GCTCGCTCCC ATGCACAACT	1500
	CTCCAAAACA CTTATCCTCC TGCAAGAGTG GGCTTTCCAG GTCTTTACT GGAAGCAGT	1560
45	TAAGCCCCCT CCTCACCCCT TCCTTTTTTC TTTCTTTACT CCTTGGCTT CAAAGGATTT	1620
	TGGAAAAGAA ACAATATGCT TTACACTCAT TTTCATTTT TAAATTGCA GGGGATACTG	1680
	AAAAATACGG CAGGTGGCCT AAGGCTGCTG TAAAGTTGAG GGGAGAGGAA ATCTTAAGAT	1740
50	TACAAGATAA AAAACGAATC CCCTAAACAA AAAGAACAAT AGAACTGGTC TTCCATTTTG	1800
	CCACCTTTCC TGTTCATGAC AGCTACTAAC CTGGAGACAG TAACATTTCA TTAACCAAAG	1860
55	AAAGTGGGTC ACCTGACCTC TGAAGAGCTG AGTACTCAGG CCACTCCAAT CACCCTACAA	1920
	GATGCCAAGG AGGTCCCAGG AAGTCCAGCT CCTTAAACTG ACGCTAGNCA ATAAACCTGG	1980
60	GCAAGTGAGG CAAGAGAAAT GAGGAAGAAT CCATCTGTGA GGTGACAGGC AAGGATGAAA	2040

	GACAAAGAAG GAAAAGAGTA TCAAAGGCAG AAAGGAGATC ATTTAGTTGG GTCTGAAAGG	2100
	AAAAGTCTTT GCTATCCGAC ATGTACTGCT AGTACCTGTA AGCATTTTAG GTCCCAGAAT	2160
5	GGAAAAAAA ATCAGCTATT GGTAAATATAA TAATGTCCTT TCCCTGGAGT CAGTTTTTTT	2220
	AAAAAGTTAA CTCTTAGTTT TTACTTGT TTCTCTAAAA GAGAAGGGAG CTGAGGCCAT	2280
10	TCCCTGTAGG AGTAAAGATA AAAGGATAGG AAAAGATTCA AAGCTCTAAT AGAGTCACAG	2340
	CTTTCCCAGG TATAAACCT AAAATTAAGA AGTACAATAA GCAGAGGTGG AAAATGATCT	2400
	AGTTCTGAT AGCTACCCAC AGAGCAAGTG ATTTATAAAT TTGAAATCCA AACTACTTTC	2460
15	TTAATATCAC TTGGTCTCC ATTTTCCCA GGACAGGAAA TATGTCCCCC CCTAACTTTC	2520
	TTGCTTCAAA AATTAAAATC CAGCATCCCA AGATCATTC TACAAGTAATT TTGCACAGAC	2580
20	ATCTCTCAC CCCAGTGCCT GTCTGGAGCT CACCCAAGGT CACCAAACAA CTGGTTGTG	2640
	AACCNACTG CCTTAACCTT CTGGGGGAGG GGGATTAGCT AGACTAGGAG ACCAGAAGTG	2700
	AATGGGAAAG GGTGAGGACT TCACAATGTT GGCCTGTCAG AGCTTGATTA GAAGCCAAGA	2760
25	CAGTGGCAGC AAAGGAAGAC TTGCCCAGG AAAAACCTGT GGGTTGTGCT AATTCTGTC	2820
	CAGAAAATAG GGTGGACAGA AGCTTGTTGG GTGCATGGAG GAATTGGGAC CTGGTTATGT	2880
30	TGTTATTCTC GGACTGTGAA TTTTGGTGAT GTAAAACAGA ATATTCTGTA AACCTAATGT	2940
	CTGTATAAAT AATGAGCGTT AACACAGTAA AATATTCAAT AAGAAGTCAA AAAAAAAAAA	3000
	AAAAAACTCG AGGGGGGGCC CGGTACCCAA TTTNCCAAAT AGAGATNGTA TTAC	3054
35		

(2) INFORMATION FOR SEQ ID NO: 195:

- 40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 907 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
- 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 195:

	GGCAGAGCTC GTGGCCGNAA CTTTTCTGCT TCCTGGCTGC CACCTACTGG CTGGCCGCGG	60
50	CCCTGGCCTG GGCTGCACC AGCTGCGNG CGGCTCCCA CAGCAGCCCC CTTCCAAGCA	120
	GCGTCCCCAC ACCGCGCACC TTCTGCGGA ACGTCTCGC CGTGCCGGGG ACCATATGGA	180
55	CGGAAGGCTT TGTGCTCACC TACAAGCTGG GTGAGCAGG TGCCAGCAGC CTGTTGATCC	240
	TCCTGGCTCC TGCTGGAGCA CGAGCGCGT TTCTGCTCCC GAGTGGGAC TGTGGAATGG	300
	TGTGGTGCT GTGGTCTGCT CCATCGCTGG CTCCTCCCTG GGTGGACCT TGCTGGCCAA	360
60	GCACTGAAAA CTGCTGCCTC TGTGAGGTCG GTGCTGCGCT TCCGCCTCGG GGGCCTAGCC	420

5 TGTCAGACTG CCTTGGTCTT CCACCTTGGA CACCTTGGG GCCAGCATGG ACGCTGGCAC 480
AATCTTGAGA GGGTCAGCCT TGCTGAGCCT ATGTCTGCAG CACTTCTTGG GARGCCTGGT 540
CACCACAGTC ACCTTCACTG GGAATGATGC GCTGCAGCCA GCTGGCCCC AGGGCCTTGC 600
AGGCCACACA CTACAGCCTT CTGGCCACGC TGGAGCTGCT GGGGAAGCTG CTGCTGGGCA 660
10 CTYTGGSCGG AGGGCCTGGC TGATGGGTTG GGGCCACATC CCTGCTTCTT GCTCCTGCTC 720
ATCCTCTCTG CCTTCCCGT TCTGTACCTG GACCTAGCAC CCAGCACCTT TCTCTGAGCT 780
GAGTGGCTGG AGTGGTCAAT AAAGCCACAT GTGCCTGTGG CCCAAAAAA AAAAAAAAAA 840
15 AAAAAAAAAA AAAAAAATG GAGGGGGGGC CCGGTACCCA AATCGCCGGA TATGATCGTA 900
ACAATC 907

20

(2) INFORMATION FOR SEQ ID NO: 196:

25

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1290 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 196:

GGCAGGAGGA GGGACAGGGA GTGGGCAAGG GGAAGAAGCA GCTTATTTGA CTAACCAGCC 60
35 CCTCTGTGGT CCACCAGCGT CTTGGCTTGG TGGGAGGGCT CTCATCAGC AGGGCCCCAG 120
KAGGGCAAGA AGAAGTGGG CAAAGCCTGG CGCTCGGCCG CGGTGCGGC AGCTTTGCMA 180
TCTGGAGCCA CGCCTCCTCC AGGCCATGCT CCTTGAACCT GGAAATGTCA ACCGGAGCCC 240
40 TTAACACCAG CCCTCCAGCA TCTAATAGAC TTGAATCTAC TCTAAACGAA TATTTAATCC 300
AACCTCAACT ACATTGTAGC TCAGTCCAAC GACTAACCTT GAAATGGGG TGTTCAGCC 360
45 TTCAGCGAGA TGGCCAAGCG GTCOCCTGGG GGCTGTGGCA GCGGGCTTAT CCTTCTCTGT 420
TGCCAACCTT GCCGTCCGAC CTCCTCCGCC CCCATGCGGT GACCCCGTCC GTGTCTGTGT 480
CTGTCCATAC GTGTGAGTCC AGCTAAAAAG ACAAACAGA ACCCGTGGGC CCAGCTCGGA 540
50 AGGTGCGTGG AGAAGGCTCC GACGTCTCCG AAGTGCAGCC CTGGGATGG CATTCGGTTG 600
TGTGCCTTAT TCCTGGAGAA TCTGTATACG GCTCGCCTAT AAGAAATATA GCCTCTTCAT 660
55 GCTGTATTAA AAGGACTTTT AAAAGCAAAA AAAAAAAAAA AAAAATCGA GGGGGGGCCC 720
GGTACCCAAT TCGCCAATA GTGAGTCGTA TTACAATTCA CTGGGCCGTC STMTTAACAA 780
60 CGTCGTGAAC TGGGAAAACC CTGGCGTTTA CCCAACTTAA TCGCCTTGCA GCACATCCCC 840

CTTTCGCCAG CTGGCGTTAA TAGCGAAAAA NGCCCGCACC CGAATCGCCC TTCCCAACAG 900
TTTGGCGCAGC CCTGAATGGC GAAATGGCAA ATTGTAAGCG TTTAATATTT TKKTTAAAAAT 960
5 TCCNCGTTWA AWTTTTTGTT TAAATCARCT CAATTTTTTT AACCCAATAA GSCCGAAATC 1020
CGGCAAAATCC CCYTTATTAA TTCCAAAAAA ATAAACCSAA AAWGGGTTTG AATTTTTTKT 1080
10 TTCCCCAYTT TTGAAACAA AWTYCCCCCT TTTTAAAAA GTTGAACCC CCAMCCYTCC 1140
AAAGGGGAAA AAACSYTTTT YTGCGGGGNA ANGGGGCCCC CNTACTTTNA ACAYCCCCC 1200
CCAAWCAATT TTTTGGGGG GTCCCNAAAG GTCCCCCTAA AANCTTTTTT CGGAACCCNA 1260
15 AGGGGANCCC CCCATTTAAA ATTTTNGGTN 1290

20 (2) INFORMATION FOR SEQ ID NO: 197:

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 1020 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 197:

30 GGTGTGCCTG GATGGTCGTG TAGGTGAGTT TTACCAAGGA TTATGGTAAC AAATGAGTGA 60
GACCTCTATG GAGAAAATAT TGAAGNNCAT TAAAGAAGAC CTCATANTAG GAGAGAATGT 120
35 SCTTTGGAGG ATTTGTATTG AGCTTTTACA GTATTCAATT TTCAACTCAA GGCAATGGCT 180
TTCTACACCA ACTCTAATCC ATAAACGGGT CTTATGACAT CTATGAAGTA GTAGCAAGAC 240
ATGCTTAGTG TGTATTTCTC TCTTTGAGAC ACTGTAATTT CTACCAGAAA TTTCCAGAGC 300
40 ATTATGTAGG TAGAAAAAAA TGCAAGCAAG CTGTTAAAGA TCTTGGATCC CATTATATAG 360
TATGTATAGC TGAAATCTGT AATTCAATCA CTTTCTCTCT TTTATCCTCT AACCAAAAAA 420
45 TTGTTTAATT TTGCATCCCA AATGTTTTTA ATCTTTGTAT ATTTTATAA AAYCCTTTTC 480
TCCTCATCAT TGCCTTTTTT GTGGTTGTAA ATAGACTTAC TTGCACTTTG AAGATGAGTT 540
ACTCCTTGTC ATCTTACAAA TATGTGATAT GGTAATTTTC ATAACAGATG TCAGTTTGA 600
50 ACCAAGAATT GGTGATTTGT TTATAAGAAA AAAACTGGCT TCATTTCTGT GAAATTGCTC 660
TTTGAAAATT TCTTTTTTACA CGTGAAGCC AACTGAGATA CCGTGATGGT GTTGATTCTT 720
TTCAATGATG CTTACCATCT ATTTTAGCCA CTGAGCCTTT TATTATTTGT CTATTGTAA 780
55 AGTTTATTTG TCTTAACTCA TTTAATAAAT ATACTGTTTA TCTGTTCTG AATGGGGACT 840
GAACTTTTTG GATATTGATA TTGATTGAA AATATTTTGG AATTTTCTCT ACTTGAAATT 900
60 TTAGAAATCT AATKGAAAAT TCTATAATGT ACTGAAAGTA WGGTTGTGTA CAGTGAKCAC 960

TCTCTAATAA TATGATGNCT TGCCCTAAAN GAGNGGGAC ATGTCCCACT TTCCACCACG 1020

5

(2) INFORMATION FOR SEQ ID NO: 198:

- 10 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 524 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 198:

AATTCGCGAA GCTGAGGGTT GTGTGCCNTC GGGCGAGCCA AGTCTTTTGA CCGGACCCTT 60
CCCGGCGCAG AAGANCTGAA GTTGATTGTA GAGCCTGTKT TTGGGGTTTA GCCGAGCTGC 120
20 TCGGGGCTTY GTGCGCGGCC AGGACACAAG YTACTTGCAA CGGGGCGGCG CCTGGCTTAT 180
GATGTTCTTC AACCCAGGGG CGGCTCTGTC CCTCTACTCG TGCCAGGCCC ACTTGCCAGG 240
25 CAGGAGCCCT CCCCAAGCCT TCAGGGCTGC TCGGAGTCAC CTGTTGGAAT GGAATAAAAG 300
GACCCCTGTG TGGGAACAGG TGCTCCAAAC ACCCTGCTGC TGGCTGCCAG GCAGGCCCTC 360
TGGAAGGGAA GGGGAGGAC TCATCAGGAC CTCCTGGAC CCTGCAGGC AGGCAGTTGG 420
30 CCCGAGCCA AGCATTTGGC TCTGCTTGCC CCAAGGGGAC AGGAAGCCTC TTGGGCTCTT 480
TCCCTTCTG GACAAGGCCC CTGCTTTTG CCTACATAA ACTG 524

35

(2) INFORMATION FOR SEQ ID NO: 199:

- 40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 332 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 199:

GTGATACAAG GAAGGTGAT CATCATCTGT CACCATGCAA TTCCTGCTCA CAGCCTTTCT 60
50 GTTGGTGCCA CTCTGGCTC TTGTGATGT CCCCATATCC CTAGGCTTCT CCCCTCCTA 120
GAAGGGCTTC TTGATAGATT AGAAAATAAG AATGAGTGAC ATTTCTTATG TGCATATAAG 180
AAGGAGCCAC AAGACATGTC TTTTAAATAA AAGGACAGTG TCCATCCTTT TAGCTGCCGA 240
55 ATAGAACCTT GGTCTCATCC TCCTGGAGCT AGGSCTTAAA ACAGCTTCTG TGTCTTSAT 300
TKGTCTCART GTTTTGCCAA GGTCTTATTC GG 332

60

(2) INFORMATION FOR SEQ ID NO: 200:

5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 376 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 200:

CCAGGGAAGC CCCARGCCTG TCCTGAATG ACATCAGTGC TTCCCTGAAC TGCCTCCCCC	60
15 ACCCTTGGGC ATTATCCCAG GAAACTTATG TTTTCTAGAA GCTAAGCAGC TGCTGGGACT	120
CAGGGACTGG TGCAGGTAGG CTGAGTGGCA GCTCAGTCCT AGAAGGTCTC TGAAGATCTG	180
GACTGAGGAC CYTGCTACTC CCCAAGCCAG AGCCCATCAG CCAGGCCTGC TGTGAGCCAC	240
20 CTGCCTGTGG AGTGCTGAGC TCAACCAAAG GCTGGCAAGC TCTGGGCCTC ATTTAAGGGA	300
TTCTGATGAG CCGATGGGCC CTGGAGGCAG CCCATTAAAG CATCTGGCTC GTTTTGGAA	360
25 AAAAAAAAAA AAAAAG	376

(2) INFORMATION FOR SEQ ID NO: 201:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1192 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 201:

40 CCCAGTATAT TTCTATAACA TTTATTTTAG TGAACCTATA ATGTTTCTTT GTATTAAATT	60
ATTAGATTAT ATCTTTAGAT AATATTGTTA CTNAATTAGT AGGTAATATA TATTTTATTC	120
AAAAATAAAT TGTGCATCTA ATGTCTACCA ATTAATGTAC TTGTAGATGT ATCTTATCTT	180
45 AACTTGAGTC TTTGCTGCCC CTAATGAGGT GTGAAGGACT CTTCTCCCCT GGGGAAGTTT	240
TTCTTTTTC A GGAGGAGGA GGGCTTTCCC AGGTAATGTG TCTAGAGTGT TGGGCAGAAR	300
50 AATCTGGGAC CACACCACAC CAGTTCTCTC CTTAATCCAC GTCATTGCCC TTCTATCCCA	360
GCTATGTTTC CAGTGTCTC TGGGTGTTTC CAAGACCAAC AAGAAAYGAA TAAATCTCTG	420
KTGAGTTGTT TATTTGTTCT TCACTTTGTT TTACACTGTA WTTTCTGAGT TTATGGGTGT	480
55 CTGTGAATTA AAAAGGAAAA GTRGAAATAA GTAAACTCA GGTGAAGGA AATATACATA	540
AATAAGATAA AGCTGACCTG TAGATATARR CAGGTTATAA RAGCTTAGAG TTGTCTAAGT	600
60 TGRGTGCAAA KTTTCTCTG ATCTTTCTGA TGCCGARACA AAAAAGGCAG TCATGTTTGT	660

5
10
15
20

WATGTGATTG GAATGGAACC CGARAAGAGA GCAYGCTGTG TTCTTGGGGA CAGGAAAGCT 720
TGYGTGCACC AAGTCTKAAC CACCACCTTC ATGGGACATA GRTTATGTGC TGGAACATAT 780
TTCACACCGG CCTGGCAGTA AACACTTGTA GTGTTGTGCA GTGGAAACGG TCATCTTCCG 840
CTAAAGCAGC GCGTGTGTG CAGCGGAAAT GGTCACTGTC TGCTAAAACA CAGCTTCCAT 900
CGTAATGTAT GCTCCTTACT CAAAGAGTGT GGTCCCAAAC AGCCTTTGGG AGGTCCCTCT 960
TGATTCATGG ATGAAACCTG GAACATCTTG AGGACTGAGT TAACCATAGG TCCTTAAATA 1020
ACTCTCCACA CGTTTTTCTT AGTTTATCTC TACATGCAGG GTGTGCAGCA GCCTGTTCAA 1080
AGTCATATTT TCTGGGAAAT ATTTCCAGTG TTTATTTGCA CTTTAGCCCA CTCTGTGTAG 1140
CCTTATTTCT TCTAAACTCA CCATTAATCT GAATAATAGT CAAATTTAGG GG 1192

(2) INFORMATION FOR SEQ ID NO: 202:

25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 589 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 202:

ATCTTGGGCT ATCTTTGACA GGGGATTCTT GCAAGTTGAT GCTTTCTACA AGTGAATATA 60
35 GTCAGTCCCC AAAGATGGAG AGCTTGAGTT CTCACAGAAT TGATGAAGAT GGAGAAAACA 120
CACAGATTGA GGATACGGAA CCCATGTCTC CAGTTCTCAA TTCTAAATTT GTTCCTGCTG 180
AAAATGATAG TATCCTGATG AATCCAGCAC AGGATGGTGA AGTACAACTG AGTCAGAATG 240
40 ATGACAAAAC AAAGGGAGAT GATACAGACA CCMGGGATGA CATTAGTATT TTAGCCACTG 300
GTTGCAAGGG CAGAGAAGAA ACGGTAGCAG AAGATGTTTG TATTGATCTC ACTTGTGATT 360
45 CGGGGAGTCA GGCAGTTCCG TCACCAGCTA CTCGATCTGA GGCACTTTCT AGTGTGTTAG 420
ATCAGGAGGA AGCTATGGAA ATTAAAGAAC ACCATCCAGA GGAGGGGTCT TCAGGGTCTG 480
AGTGGAAGA AATCCCTGAG ACACCTTGTC AAAGTCAAGG AGAGGAACTC AAAGAAGAAA 540
50 ATATGGAGAG TGTTCGTTG CACCTTTCTC TGACTGAAAC TCAGTCCCA 589

55

(2) INFORMATION FOR SEQ ID NO: 203:

60 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 847 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 203:

5
GGCACGAGCG CAAGCTGCTG GCCGCCATCA ACGCGTTCCG CCAGGTGCGG CTGAAACACC 60
GGAAGCTCCG GGAACAAGTG AACTCCATGG TGGACATCTC CAAGATGCAC ATGATCCTGT 120
10 ATGACCTGCA GCAGAATCTG AGCAGCTCAC ACCGGGCCCT GGAGAAACAG ATTGACACGC 180
TGGCGGGGAA GCTGGATGCC CTGACTGAGC TGCTTAGCAC TGCCCTGGGG CCGAGCAGCT 240
TCCAGAACCC AGCCAGCAGT CCAAGTAGCT GGACCCACGA GGAGGAACCA GGCTACTTTC 300
15 CCCAGTACTG AGTGGTGGAC ATCGTCTCTG CCACTCTCTGA CCAGCCTGAA CAAAGCACCT 360
CAAGTGCAAG GACCAAAGGG GGCCTGGCTT GGATGGGTG GCTTGCTGAT GGCTGCTGGA 420
20 GGGGACGCTG GCTAAAGTGG GGAGGCCTTG GCCCACCTGA GGGCCAGGT GGAACATGG 480
TCACCCCCAC TCTGCATACC CTCATCAAAA AACTCTCAC TATGCTGCTA TGGACGACCT 540
CCAGCTCTCA GTTACAAGTG CAGGCGACTG GAGGCAGGAC TCTTGGGTCC CTGGGAAAGA 600
25 GGGTACTAGG GGCCCGGATC CAGGATCTG GGAGGCTTCA GTTACCGCTG GCCGAGCTGA 660
AGAACTGGGT ATGAGGCTGG GCGGGGCTG GAGGTGGCG CCCCTGGTGG GACAACAAAG 720
30 AGGACACCAT TTTTCCAGAG CTGCAGAGAG CACCTGGTGG GGAGGAAGAA GTGTAACTCA 780
CCAGCCTCTG CTCTTATCTT TGTAAATAAT GTTAAAGCCA GAAAAAAAAA AAAAAAAAAA 840
AAAAAAAA 847
35

(2) INFORMATION FOR SEQ ID NO: 204:

40
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 852 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 204:

50 ACAAACATAC TCGCAGGAAG GAGTCTCATG CTGCCCCGAG CATCAGCGCA ACNCNTGGCC 60
GCCATCAACG CGTTCCGCCA GGTCCGCTG AAACACCGGA AGCTCCGGGA ACAAGTGAAC 120
TCCATGGTGG ACATCTCCAA GATGCACATG ATCCTGTATG ACCTGCAGCA GAATCTGAGC 180
55 AGCTCACACC GGGCCCTGGA GAAACAGATT GACACGCTGG CGGGGAAGCT GGATGCCCTG 240
ACTGAGCTGC TTAGCACTGC CCTGGGGCCG AGGCAGCTTC CAGAACCCAG CCAGCAGTCC 300
AAGTAGCTGG ACCCACGNAG GAGGAACCG GCTACTTTCC CCAGTACTGA GGTGGTGGAC 360
60

ATNCGTCTCT TGCCACTCCN TGNACCCAGC CCTGAACAAA GCACCTCAAG TGCAAGGACC 420
AAAGGGGGCC CTGGCTTGGG GTGGGTGGC TTGCTGATGG CTGCTGGAGG GGACGCTGGC 480
5 TAAAGTGGGK AGGCCTTGGC CCACCTGAGG CCCCAGGTGG GAACATGGTC ACCCCCACTC 540
TGCATACCCT CATCAAAAAC ACTCTCACTA TGCTGCTATG GACGACCTCC AGCTCTCAGT 600
TACAAGTGCA GGCAGCTGGA GGCAGGACTC CTGGGTCCCT GGGAAAGAGG GTACTAGGGG 660
10 CCCGATCCA GGATTCTGGG AGGCTTCAGT TACCGCTGGC CGAGCTGAAG AACTGGGTAT 720
GAGGCTGGGG CGGGGCTGGA GGTGGCGCCC CCTGGTGGGA CAACAAGAG GACACCATTT 780
15 TTCCAGAGCT GCAGAGAGCA CCTGGTGGGG AGGAAGAAGT GTAACCTACC AGCCTCTGCT 840
CTTATCTTTG TA 852

20

(2) INFORMATION FOR SEQ ID NO: 205:

(i) SEQUENCE CHARACTERISTICS:
25 (A) LENGTH: 1354 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 205:

GATTGGGCAC GAGGCTTGCT GGAGCAGGAG AAGTCTCTRG CCGGCTGGGC ACTGGTGCTG 60
GCASGARCTG GCATTGGACT CATGGTGCTG CATGCAGAGA TGCTGTGGTT CGGGGGGTGC 120
35 TCGGCTGTCA ATGCCACTGG GCACCTTICA GACACACTIT GGCTGATCCC CATCACATTC 180
CTGACCATCG GCTATGGTGA CGTGGTGCCG GGCACCATGT GGGGCAAGAT CGTYTGCCCTG 240
40 TGCACTGGAG TCATGGGTGT CTGCTGCACA GCCCTGCTGG TGGCCGTGGT GGCCCGGAAG 300
CTGGAGTTTA ACAAGGCAGA GAAGCACGTG CACAACTICA TGATGGATAT CCAGTATACC 360
AAAGAGATGA AGGAGTCCGC TGCCCGAGTG CTACAAGAAG CCTGGATGTT CTACAAACAT 420
45 ACTCGCAGGA AGGAGTCTCA TGCTGCCCGC AGGCATCAGC GCAANCTGCT GGCCGCCATC 480
AACGCGTICC GCCAGGTGCG GCTGAAACAC CGGAAGCTCC GGGAAACAAGT GAACTCCATG 540
50 GTGGACATCT CCAAGATGCA CATGATCCTG TATGACCTGC AGCAGAATCT GAGCAGCTCA 600
CACCAGGGCCC TGGAGAAACA GATTGACACG CTGGCGGGGA AGCTGGATGC CCTGACTGAG 660
CTGCTTAGCA CTGCCCTGGG GCCGAGGAG CTTCCAGAAC CCAGCCAGCA GTCCAAGTAG 720
55 CTGGACCCAC GAGGAGGAAC CAGGCTACTT TCCCCAGTAC TGAGGTGGTG GACATCGTCT 780
CTGCCACTCC TGANCCAGC CCTGAACAAA GCACCTCAAG TGCAAGGACC AAAGGGGGCC 840
60 CTGGCTTGGG GTGGGTGGC TTGCTGATGG CTGCTGGAGG GGACGCTGGC TAAAGTGGGK 900

AGGCCTTGGC CCACCTGAGG CCCAGGTGG GAACATGGTC ACCCCCACTC TGCATACCCCT 960
CATCAAAAAC ACTCTCACTA TGCTGCTATG GACGACCTCC AGCTCTCAGT TACAAGTGCA 1020
GGCGACTGGA GGCAGGACTC YTGGGTCCCT GGGAAAGAGG GYACTAGGGG CCCGGATCCA 1080
GGATCTCTGG AGGCTTCAGT TACCGCTGGC CGAGCTGAAG AACTGGGTAT GAGGCTGGGG 1140
CGGGGCTGGA GGTGGCGCCC CCTGGTGGGA CAACAAAGAG GACACCATTT TTCCAGAGCT 1200
GCAGAGAGCA CCTGGTGGGG AGGAAGAAGT GTAACTCACC AGCCTCTGCT CTTATCTTTG 1260
TAATAAATGT TAAAGCCAGA AAAAAATAAA AAAAAAAAAA AAAAACTCG AGGGGGGCCC 1320
AGACCCAATC TCCCTATAGT AAGNCGCCNN ANAN 1354

20

(2) INFORMATION FOR SEQ ID NO: 206:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1378 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 206:

TCCCCAGGTG CACAGCCAGG GCCCTCCTGT CTGCAGGAGA ATTCACAGCT GGTGTGGGAC 60
TCAGCCCCCTA GNCCATTCAA AGCCTTAATG TTGTAATCAT ATCTTACGTG TTGAAGACCT 120
GACTGGAGAA ACAAATGTG CAATAACGYG AATTTTATCT TAGAGATCTG TGCAGCCTAT 180
TTCTGTCAACA AAAGTTATAT TGTCTAATAA GAGAAGTCTT AATGGCCTCT GTGAATAATG 240
TAACTCCAGT TACACGGTGA CTTTAAATAG CATACAGTGA TTTGATGAAA GGACGTCAAA 300
CAATGTGGCG ATGTCGTGGA AAGTTATCTT TCCCGCTCTT TGCTGTGGTC ATTGTGTCTT 360
GCAGAAAGGA TGGCCCTGAT GCAGCAGCAG CGCCAGCTGT ANATAAAAAA TAATTCACAC 420
TATCAGACTA GCAAGGCACT AGAACTGGAA AAGACCACAG AAAACAAAGA ATCCAACCCCT 480
TTCATCTTAC AGGTGAACAA ACTGTGATGA TGCACATGTA TGTGTTTGT AAGCTGTGAG 540
CACCGTAACA AAATGTAAAT TTGCCATTAT TAGGAAGTGC TGGTGGCAGT GAAGAAGCAC 600
CCAGGCCACT TGACTCCAG TCTGGTGGCC TGTCTACACC AGACAACACA GGAGCTGGGT 660
CAGATTCCCC TCAGCTGCCT AACAAAGTTC CTCGAACAGA AAGTGCTTAC AAAGCTGCCT 720
TCTCGGATAC TGAAAGGTCG AGTTTCTGA ACTGCACTGA TTTTATTGCA GTTGAAAAAA 780
AAAAAAGCT ATTCAAAGA TTTCAAGCTG TTTCTGAGACA TCTTCTGATG GCTTTACTTC 840
CTGAGAGGCA ATGTTTTTAC TTTATGCATA ATTCATTGTT GCCAAGGAAT AAAGTGAAGA 900

AACAGCACCT TTTAATATAT AGGTCTCTCT GGAAGAGACC TAAATTAGAA AGAGAAAAC 960
GTGACAATTT TCATATCTTC ATTCTTAAAA AACACTAATC TTAAC TAACA AAAGTTCTTT 1020
5 TGAGAATAAG TTACACACAA TGGCCACAGC AGTTTGTCCT TAATAGTATA GTGCCTATAC 1080
TCATGTAATC GGT TACTCAC TACTGCCTTT AAAAAAAAA ACCAGCATAT TTATTGAAAA 1140
CATGAGACAG GATTATAGTG CCTTAACCGA TATATTTTGT GACTTAAAA ATACATTTAA 1200
10 AACTGCTCTT CTGCTCTAGT ACCATGCTTA GTGCAATGA TTATTTCTAT GTACAACTGA 1260
TGCTTGTTCT TATTTTAATA AATTTATCAG AGTGAAAAA AAAAAAAAAA AAAAAAAAAA 1320
15 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAGAA NAAAAANA 1378

20 (2) INFORMATION FOR SEQ ID NO: 207:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1166 base pairs
(B) TYPE: nucleic acid
25 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 207:

30 AANCCACTGC ANTTTAAACC CCTCCCTC CAAGAAAGTT CACAACCGGC CATGGATGAC 60
CCTCATTTTA GATGGCCNC AATATTTAAG ATGGACTGRG GMCCCARAG ACTGACCCTT 120
GAAAGGGGA CTCAGAAGAA AGATCCTTGA CATTGCCMAA CATGCTGGGC TTGTCCAACA 180
35 CAGTGATGCG GCTCATCGAG AARCGGCTT TCCMAGGACA AGTACTTTAT GATAGGTGGG 240
ATGCTGCTGA CCTGTGTGGT CATGTTCTC GTGGTGCACT ACCTGACATG AGCCAGCCAC 300
40 GCTCAGTGGC TGAACAGCAT TCCACAGCC TGCAAGTGTG TGTGTGTGTG AAAGAGAGAG 360
GGGGCCAGA GCGCGCTTT TGAAATGTTT GCCTGTCTGA ACTGTGAAGA CACTTGGGAG 420
TGATGTGGT CTAATTTC AACTGCTCTG TTTCTGTGA CATCTTGGAG GGGAGCTAG 480
45 TGCCAMCACC ATGCGCGTG CTAGGAAAT GAAAGAAGTC CCGGTCTGT CTCTCTCACT 540
CTCGCTCTCA MTGGGGGAGG GAAAGAATGG CTTGGTGGC TTTGTTTACA CAGCTGATGC 600
50 GTGSCTGGG AAGGTGTCCA CAGTGAGCCC TGTGTGCAGG ACTGTCCACN ACGTTTACA 660
CCTTGTACC ATCAGGCTT TCTGGCTCCT GATAGGTGG AGCAAAAGTG GAAAGGAAAG 720
GAAAGAGGCY TTTTCTACA GCCATTATAT TAAATAGTAG GTCGATTAC ATCYTCGTGC 780
55 TCCTGGCCAC CCTCCCTGT GCCTCAGTGA CATGTAGATG ACTGACTGCC AATACTTGTG 840
ACCATTCCT GGAAGCAGCT ACCTAGGGA AACAAGATGT AGTGCTATTG CCGATAACAA 900
60 GTAAGATTTT CCACACTACA GCTGGGTGTT TCTCTTTTCT AAAGTGAGGC CAGTGTATT 960

5 TCCCGGGAGT GTTCAGTCTT GACCCTAGTC ACTGATTTT TCTAGTTGTT AATAGAGTGG 1020
TTGGGCTTTT AAGGTTT CAGA GACTGTGGGC TTGGGCACCT GCGCCCAGGG STTTTGTGGG 1080
GGCCTTTGCC CCTTAGRAAA GTAGCTTTTA GGGGCAAAGA TTTGTGATT TTCCCCATTA 1140
CAGTCTTCAG CTCNAGGGTT TTAATA 1166

10

(2) INFORMATION FOR SEQ ID NO: 208:

15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 697 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 208:

TACTTCTAGG ATTATAAGGA ATTAACATTG AGATGACATT TCCATTGAG AAGGAAAATA 60
25 GTTGCTTTCA GTGCCTTTTA TTTGATTCTT GGAGAGAGCA GACTCGCACS AACATTCAAC 120
CCCAGCGCTG ATATGACAGT AATCCTCAGA GGCAGAGCCC AGCACAAAAC AGCAATGCTA 180
GAAAGTTACA ATTGGAAAGT TTCCTGCCAG CTTCGGGAAT GACACTGCAA AGCTGATGCC 240
30 AGAAACTGCC AGRGTAATTC TCCTCATTAC TGCTCTACCC ACCCACTTTC AGCTCCCCAA 300
ATTAAGTAGT GCAGTTGACT AATTCTCTTT ACCTTTATCA TTTARGGTGA RGCATTGCAC 360
35 AAAAACTCTC GACTTTGCCA TATAAGGGCT GTGGTTCTCT GTGGTCCCCT GGATAAGAGG 420
CATCACCATT ATCTGGAAAC ATGCAGTAAA TGCAGATTNT TCATCTTCTC CCCAGACCTC 480
CTGAGTTAGA AATTACACAAG TTCTCCAGGT GATCTCATAC ATGCTAAAAGT TTGAGAACCA 540
40 TTGAGTAAAG TTAATGCATT AAGAAGAGAT TAGATAGGGA TGGTGGCGTA TCTTCCTACA 600
GTTTCCTGT TAACAAGAAA GTCAGAGGTC AGTTGATCAG ACATTAGATT ATTTATTGCT 660
45 AAAACTAAAA AAAATTAAAA AAACTGGAG GGGGGCC 697

50 (2) INFORMATION FOR SEQ ID NO: 209:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 932 base pairs
(B) TYPE: nucleic acid
55 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 209:

60 CGTGAGTCAC CTCTCTATAG TGGGCGTGGC CGAGGCCGGG GTGACCCTGC CGAAGCCTCC 60

5 GCTGCCAGAA ACCATGTTCA AGGTAATTAA AAGGTCCTG GGGCCAGCCA GCCTGAGCTT 120
GCTCACCTTC AAAGTCTATG CAGCACCAAA AAAGGACTCA CCTCCCAAAA ATTCCGTGAA 180
GGTTGATGAG CTTTCACTCT ACTCAGTTCC TGAGGGTCAA TCGAAGTATG TGGAGGAGGC 240
AAGGAGCCAG CTTGAAGAAA GCATCTCACA GCTCCGACAC TATTGCGAGC CATAACAAC 300
10 CTGGTGTGAG GAAACGTACT CCCAACTAA GCCCAAGATG CAAAGTTTGG TTCAATGGGG 360
GTTAGACAGC TATGACTATC TCCAAAATGC ACCTCCTGGA TTTTTCCTGA GACTTGGTGT 420
TATTGGTTTT GCTGGCCTTA TTGGACTCCT TTTGGCTAGA GGTTCAAAA TAAAGAAGCT 480
15 AGTGATATCG CCTGGTTTCA TGGGATTAGC TGCCTCCCTC TATTATCCAC AACAAGCCAT 540
CGTGTTCGCC CAGGTCAGTG GGGAGAGATT ATATGACTGG GGTTCACGAG GATATATAGT 600
20 CATAGAAGAT TTGTGGAAGG AGAACTTTCA AAAGCCAGGA AATGTGAAGA ATTCACCTGG 660
AACTAAGTAG AAACTYCAT GYTCTGCCAT CTTAATCAGT TATRGGTAAA CATTTGGAAC 720
TCCATAGAAT AAATCAGTAT TTCTACAGAA AAATGGCATA GAAGTCAGTA TTGAATGTAT 780
25 TAAATTGGCT TTCTTCTTCA GGAAAACTA GACCAGACCT CTGTTATCTT CTGTGAAATC 840
ATCTTACAAG CAACTAACC TGAATCCCT TCACCTAGAG ATAATGTACA AGCCTTAGAA 900
30 CTCCTCATTC TCATGTTGCT ATTTATGTAC CT 932

35 (2) INFORMATION FOR SEQ ID NO: 210:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 661 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 210:

45 GTCATTCTTT AAATAAAGC TTTCTGTTT AAAGCTTTTC AAAGGAGCAG ACCACCTTGA 60
AGATTCCTCC TAGGGTTGAT ATGTGTCTAA TTCATTTTAT AAAAATTATT CTGTCTTCA 120
TTTTAAAGCT TTGGCTATAT AGTCAGAAAT GTCCTAAATA ACAAACTATT TTGTATTTAA 180
50 TTTAGGGAAG ACTAAAGGA AGAAAAATGA AAAGTCAGTC TTTATGTAAG CTCCAAGGAT 240
ATTAGGGCTT AAAGGCTTT TCTAGTTTTA TGAGAATTTG TACTACTGAT TTTTATATAT 300
55 TCCGTGTTTT GAGATGAACA GATCTCTGGG GAAATTTGTT AGTTACAATG GCATTTCACT 360
GTGATCCCTC TCAAGCTCAG ATCAGTTCTA TAACCCAATG ACAACCTGTC TCTTTGGTTT 420
ACTGTCCTGT GAAATGTCAG CTCAAGTTTC CCAGAAGTCG TGTGTTTATG ATGAGTCAGA 480
60

GTGCTTTTCC TCGGTGGGAC AGTTGCTGGC CCTCTTAATT TTGGTGTATG TGCTTCCAAG 540
TATCTAAACC TCCAGTCTGA TCTGTATATG CTATCCTAAC TGTTAATTGT ATTATTGATT 600
5 ATGTTGATTA TCTTGCTTGA AGGTTTCATAC TTTTCAATTT GATAGAAATA AAGTTTTTTT 660
C 661

10

(2) INFORMATION FOR SEQ ID NO: 211:

(i) SEQUENCE CHARACTERISTICS:
15 (A) LENGTH: 592 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 211:

GAAACTGACA TTGTTAAACA CACTAAAACA GAAGTACTTA CCTCTTGAAG ATTTAATATA 60
TAATGGTTGA CATGATACAT GTACATGAAT GGAATGACCA GATGCTTATG GTCTACATTT 120
25 TCCTTTATCC TGTTAGTATT ACCTTCCTTA ATCTTTGTTT CTTAACATGC TAAATTCCCTC 180
TTCAGTGTTC ATTTTCTAGT GACAGAATGC TAACATTTCT TACACCCTGG CAGAAGGGAG 240
30 AGAAATGTGT TTTGGGTGG GTAACATAAT TTTTGAGTGA AATATCATAA GATGAGAATG 300
GAAAGAGGGA GACACAAAGA GTTATAACAA AAAACAATG GTTTTTTTAG CCAATTGACT 360
GGCTCTTTAA ATAGTCTACA AGACATTCAC GTTNAACATC ACTTTTAGTG AAATAAAATG 420
35 TGCCATACTA GTATGTGCTT CAAAAGGGCA AATGTGCTTT AGTGCCCTAA GGCTAAATTT 480
TGGTCATTTG ACATCAGAGA TGTGTAAAGT ATTGCACCTA ATACGCACCT ATTTCTCAAT 540
40 AGTGNATTTT TTTTGGCTAG CATTNCTTTT ACCACTAACC TTGTTGGATA GC 592

45 (2) INFORMATION FOR SEQ ID NO: 212:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 938 base pairs
(B) TYPE: nucleic acid
50 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 212:

55 TGGAGTGGCT TTCCAGCTGA ATGAATCCTA TGTCTCGCGT GCAGGTGGTT GGTTTTCAAT 60
GTCTTTSCTA ATTTTTTTCC TATGGCTCT TGGGAGTTT CTTTGTTTGC TCCTGTGTTT 120
GCCAGCTTT AATAAAACCA GCGCAAACA AAAACCATAG CATTCTGAAA CAATAGGGGG 180
60

CCCACATGG ACCCAGTATG TCACTTTAAT GGACTTCAAG AAAAAATCTG AATGGGAAAA 240
TGACACTAGG AATGTATACT CCACACATTT TATGCCATAT AATGGTGTGT TTTCTTAATT 300
5 TTGTTTCTTG TGGCGAAATG TGGCTTTCAA ATTAAATGM CCTTTTCTTC TTKGAACTT 360
TTTGTMTKGA CTKGATATAAT TAAGGGTTTG GAAAGATTCA TAATMTGAG AGAGGTTTGC 420
AACCAGGAGA TACAAAGAAG TCTCAGTAGT AATCTTGTTT ATGTGCTTTT ACAGCCAGCT 480
10 ACATTTAAGR ATGTATTAGT TACAGAAATT ATATGTCTGT GTATGTGTCT CTAACAATA 540
AAGTACATGC CTCCACATAA TCGGTGCTG TCCATCTCGG CAAATACTGG CCAAGTCCCT 600
15 TTATGACAGG CACACAGAAA CCATAGCATG GTCTGGCTTT CAGAAAATGC CTCTCATCTT 660
TCCTGGAACC TTATTTTGCT AAATGTCTGT TTTCTGTGA TTTGTGTAC CTCACAGCAC 720
CATGTGACC ATGGTGATGC CTCATTTGCA TGATATGTAC CTGTGTGTTA ATGTGAAATA 780
20 CATTTTCATT GAAGAGTCTG ATGACTTGCT AGCGTTTAT TTTTCTGTA AGCTCAATGT 840
GCTGAAACCA AACCAGGCTT TTA AAAACCT GTGTAGAAGA AAACCAAAAA ATCCTGTGTG 900
25 GGTGTCCTTT CCTGTCAA CTCATTAAAA ATTCCTTT 938

30 (2) INFORMATION FOR SEQ ID NO: 213:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 1079 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 213:

40 AGCCTGCCGG GAGAGTGGT GCATCTRARA GGCTGGTCGT GGACTGTGGT TGGGGGAGGT 60
GGGAGCTGTT TTAACCGTGT GCCCCTCTC CTGTGCKGC GTGGGCATCC CCCGGGGCAG 120
TGAACGCGG GCGTCTCC AGCTTCCGAG TCCAGCCAGC CTGGGCGCGG GCGCGCCCC 180
45 CGAGACACCC GAGGAGTCCG TTCCTCCCTG GTTACGTGA CTGTGGAGCT GGTCTCTTGT 240
GGCTCAGCGC CGTGCGGAGG TTGAAGCGTA CCTGCGGAGG TCGCACCAGG GCGGTGAGGA 300
50 GGAGGAGGAA GGGCATGAGC CGAGCTTGAG GAATCCGTGY TCCAACTCT AACTCAAGG 360
RTGCMCTGCG CAACTCTGGT GCGATGGGC TGGGGCAGAT GTCCTTGAG TTCTACCAGA 420
AGAAGAAGTC TCGCTGGCCA TTCTCAGACG AGTGATCCC ATGGGAAGTG TGGACGGTCA 480
55 AGGTGCATGT GGTAGCCCTG GCCACGGAGC AGGAGCGGCA GATCTGCCGG GAGAAGGTGG 540
GTGAGAACT CTGCGAGAAG ATCATCAACA TCGTGAGGT GATGAATCGG CATGAGTACT 600
60 TGCCCAAGAT GCCCACACAG TCGGAGGTGG ATAACGTGTT TGACACAGGC TTGCGGGACG 660

5 TGCAGCCCTA CCTGTACAAG ATCTCCTTCC AGATCACTGA TGCCCTGGGC ACCTCAGTCA 720
CCACCACCAT GCGCAGGCTC ATCAAAGACA CCTTGGCCCT CTGAGCGTCG CTGGATCTCT 780
GGGAGCTCCT TGATGGCTCC CAGACCTTGG CTTTGGGAA TTGCACTTTT GGGCCTTTGG 840
GCTCTGGAAC CTGCTCTGGG TCATTGGTGA GACTTGGAAG GGGCAGCCCC CGCTGGCTTC 900
10 TTGGTTTGT GGTGGCCAGC CTCAGGTCAT CCTTTTAATC TTTGCTGACG GTTCAGTCCT 960
GCCTCTACTG TCTCTCCATA GCCCTGGTGG GGTCCCCCTT CTTTCTCCAC TGTACAGAAG 1020
AGCCACCACT GGGATGGGGA ATAAAGTTGA GAACATGAGT TTGGGCTGAA AAAAAAAAAA 1079
15

(2) INFORMATION FOR SEQ ID NO: 214:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3791 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

25

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 214:

30 TGAAGCAGGC GCTCTTGGCT CGGCGCGGCC CGCTGCAATC CGTGGAGGAA CGCGCCGCCG 60
AGCCACCATC ATGCCTGGGC ACTTACAGGA AGGCTTCGGC TCGTGGTCA CCAACCGATT 120
CGACCAGTTA TTTGACGACG AATCGGACCC CTTGAGGTG CTGAAGGCAG CAGAGAACAA 180
35 GAAAAAAGAA GCCGGCGGGG GCGGCGTTGG GGGCCCTGGG GCCAAGAGCG CATCAGGGCC 240
GCGGCCCAGA CCAACTCCAA CGCGGCAGGC AAACAGCTGC GCAAGGAGTC CCAGAAAGAC 300
CGCAAGAACC CGCTGCCCCC CAGCGTTGGC GTGGTTGACA AGAAAGAGGA GACGCAGCCG 360
40 CCCGTGGCGC TTTAAGAAAG AAGGAATAAG ACGAGTTGGA AGAAGACCTG ATCAACAAC 420
TCAGGGTGAA GGGAAATAA TTGATAGAAG ACCAGAAAGG CGACCACCTC GTGAACGAAG 480
45 ATTCGAAAAG CCACTTGAAG AAAAGGGTGA AGGAGGCGAA TTTTCAGTTG ATAGACCGAT 540
TATTGACCGA CCTATTGAG GTCGTGGTGG TCTTGAAGA GGTGAGGGG GCCGTGGACG 600
TGGAATGGGC CGAGGAGATG GATTTGATTC TCGTGGCAA CGTGAATTG ATAGGCATAG 660
50 TGGAAGTGAT AGATCTTCTT TTTACATTA CAGTGGCCTG AAGCACGAGG ACAAACGTGG 720
AGGTAGCGGA TCTCACAAC TGGGAACGT CAAAGACGAA TTAACGACT TGGATCAATC 780
55 AAATGTGACT GAGGAAACAC CTGAAGGTGA AGAACATCAT CCAGTGGCAG ACACTGAAAA 840
TAAGGAGAAT GAAGTTGAAG AGGTAAAAGA GGAGGGTCCA AAAGAGATGA CTTTGGATGA 900
GTGAAGGCT ATTCAAAATA AGGACCGGGC AAAAGTAGAA TTAAATATCC GAAAACCAA 960
60

	TGAAGGTGCT GATGGGCAGT GGAAGAAGGG ATTTGTTCTT CATAAATCAA AGAGTGAAGA	1020
	GGCTCATGCT GAAGATTCGG TTATGGACCA TCATTTCCGG AAGCCAGCAA ATGATATAAC	1080
5	GTCTCAGCTG GAGATCAATT TTGGAGACCT TGGCCGCCA GGACGTGGCG GCAGGGGAGG	1140
	ACGAGGTGGA CGTGGGCGTG GTGGGCGCCC AAACCGTGGC AGCAGGACCG ACAAGTCAAG	1200
10	TGCTTCTGCT CCTGATGTGG ATGACCCAGA GGCATTCCCA GCTCTGGCTT AACTGGATGC	1260
	CATAAGACAA CCTTGGTTCC TTGTGTAACC CTTCTGTTC AAGCTTTTGC ATGCTTAAGG	1320
	ATTCCAACG ACTAAGAAAT TAAAAAATAA AAGACTGTCA TTCATACCAT TCACACCTAA	1380
15	AGACTGAATT TTATCTGTTT TAAAAATGAA CTTCTCCCGC TACACAGAAG TAACAAATAT	1440
	GGTAGTCAGT TTGTATTTA GAAATGTATT GGTAGCAGGG ATGTTTTCAT AATTTTCAGA	1500
20	GATTATGCAT TCTTCATGAA TACTTTTGTA TTGCTGCTTG CAAATATGCA TTTCCAAACT	1560
	TGAAATATAG GTGTGAACAG TGTGTACCAG TTFAAGCTT TCACTTCATT TGTGTTTTTT	1620
	AATTAAGGAT TTAGAAGTTC CCCCAATTAC AAAGTGGTTT TAAATATGG ACATACTGGT	1680
25	TTTAATACCT GCTTTGCATA TTCACACATG GTCAACTGGG ACATGTTAAA CTTTGATTTG	1740
	TCAAATTTTA TGCTGTGTGG AATACTAACT ATATGTATTT TAACTTAGTT TTAATATTTT	1800
30	CATTTTGGG GAAAAATCTT TTTTCACTTC TCATGATAGC TGTATATAT ATATGCTAAA	1860
	TCTTTATATA CAGAAATATC AGTACTTGAA CAAATTCAAA GCACATTTGG TTTATTAACC	1920
	CTTGCTCCTT GCATGGCTCA TTAGGTTCAA ATTATAACTG ATTTACATTT TCAGCTATAT	1980
35	TTACTTTTTA AATGCTTGAG TTTCCCATTT TAAATCTAA ACTAGACATC TTAATTGGTG	2040
	AAAGTTGTTT AAAGTACTTA TTGTTGGTAG GCACATCGTG TCAAGTGAAG TAGTTTTATA	2100
40	GGTATGGGTT TTTTCTCCCC CTTCAACAGG GTGGGTGGAA TAAGTTGATT TGGCCAATGT	2160
	GTAATATTTA AACTGTCTG TAAATAAGT GTCTGGCCAT TTGGTATGAT TTCTGTGTGT	2220
	GAAAGTCCC AAAATCAAAA TGGTACATCC ATAATCAGCC ACCATTTAAC CCTTCCTTGT	2280
45	TCTAAAACAA AAACCAAAGG GCGCTGGTTG GTAGGGTGAG GTGGGGGAGT ATTTTAATTT	2340
	TTGGAATTTG GGAAGCAGAC AGCTTTACTT TGTAAGGTTG GAACAGCAGC ACTATACATG	2400
50	AAATATAAAC CAAAAACCTT TACTGTTTCT AAATTTCTTA GATTGCTATT ATTTGGTTGT	2460
	AAGTTGAGTA TTCCACAGAA AGTGGTAATT ATCTCTTCTC TCTTCCTCCA TTAGAAAATT	2520
	AGGTAAATAA TGGATTCTTA TAATGGGAGC ATCACCACCT ATTAACACAC ACATAGAATG	2580
55	ATGAATTAAA AAAGTTTCT AGGATTGTCT TTTATTCTGC CACATTTATT GATAAACAGT	2640
	GAAGGAATTT TTAATAAATT TTAAAGAATT GTTGTGACG TCATTTTAG AAATGTTCTA	2700
60	CCTGTATATG GTAATGTCCA GTTTTAAAAA TATTGGACAT CTTCAATCTT AAACATTTCT	2760

	ATTTAGCTGA TTGGTTCTCA CATATACTTC TAAAAGAAAC TTTTATGTTA TAAGAGTTAC	2820
	TTTTTGGATA AGATTATTA ATCTCAGTTA CCTACTATTC TGACATTTTA GGAAGGAGGT	2880
5	AATGTGTTTT AATGATGGAT AAACCTGTGC TGGTGTTTTG GATCTTATGA TGCTGAGCAT	2940
	GTTCCTGCACT GGTGCTAATG TCTAATATAA TTTTATATTT ACACACATAC GTGCTACCCA	3000
10	GAGATTAATT TAGTCCATAT GAACTATTGA CCCATGTGTC ATTGAGACAG CAACATACGC	3060
	ACTCCTAAAT CAGTGTGTTT AGACTTTTCA AGTATCTAAC TCATTTCCAA ACATGTACCA	3120
	TGTTTTATAA ACCTCTTGAT TTCCAGCAAC ATACTATAGA AAACACCTGC TACTCAAAAC	3180
15	ACAACTTCTC AGTGTCAATC ATTGCTGTGC TGAGAGACAA CATAGCAATA TCTGGTATGT	3240
	TGCAAGCTTT CAAGATAGCC TGAACTTAAA AAGTTGGTGC ATTAGTTGTA TCTGATGGAT	3300
20	ATAAATTGTC CTCCTAGTTC ACTTTGTGTC AAGAGCTAAA ACTGTGAACC TAACTTTCTC	3360
	TTATTGGTGG GTAATAACTG AAAATAAAGA TTTATTTTCA TGCTCACTTC TTAAAAGTCA	3420
	TAAAAACAAT CAAATAGGRT CATGTTTATT GTCATGTGTT TCCTGGKTTT TGACCTGTGT	3480
25	GCACACCCCT GTGTGTTTAT AATTTTAAAT TTGAATTTTA TATGGGGTTT TTATTTGCTA	3540
	AAAACCAGGC TGTGTAATCA CATTGGGAA GGGTACTTAT CTTAATGACT AATGACTTAA	3600
30	TTGGGAAAGT TGAATTCTTG TAAAATACAA AATCCAAGGA CTTCTTGGGA TTTAATCTAA	3660
	TTGTCACTTC NTTAGGCAGA TNCACTTTTT TGGATAATGG AAAGTTAAGC ATACCGAATG	3720
	CTACTTTTGG TTGACAAACG GGCCTAATAG TCCGGGGGGA AATCCCTAAC NGTAAGGNT	3780
35	CCCAAGTATG G	3791

40 (2) INFORMATION FOR SEQ ID NO: 215:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 1334 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 215:

50	CAGTGCTCGC TCCTGCTCGG GCGCTGCGG CCCCAGGCGT CGCCATGACC AGTGAGCTGG	60
	ACATCTTCGT GGGGAACAGA CCCTTATCGA CGAGGACGTG TATCGCCTCT GGCTCGATGG	120
55	TTACTCGGTG ACCGACGCGG TGGCCCTGCG GGTGCGCTCG GGAATCCTGG AGCAGACTGG	180
	CGCCACGGCA GCGGTGCTNC AGAGCGACAC CATGGACCAT TACCGCACCT TCCACATGCT	240
	CGAGCGGCTG CTGCATGCGC CGCCAAGCT ACTGCACCAG YTCATCTTCC AGATTCCGCC	300
60	CTCCCGGCAG GCACTACTCA TCGAGAGGTA CTATGCCTTT RATGAGGCCT TTGTTCCGGA	360

5 GGTGCTGGGC AAGAAGCTGT CCAAAGGCAC CAAGAAAGAC CTGGATGACA TCAGCACCAA 420
AACAGGCATC ACCCTCAAGA GCTGCCGGAG ACAGTTTGAC AACTTTAAAC GGGTCTTCAA 480
GGTGGTAGAG GAAATGCGGG GCTCCCTGGT GGACAATATT CAGCAACACT TCCTCCTCTC 540
TGACCGGTTG GCCAGGGACT ATGCAGCCAT CGTCTTCTTT GCTAACAACC GCTTTGAGAC 600
10 AGGGAAGAAA AAAGTGCAGT ATCTGAGCTT CGGTGACTTT GCCTTCTGCG CTGAGCTCAT 660
GATCCAAAAC TGGACCCCTG GAGCCGTCGA CTCACAGATG GATGACATGG ACATGGACTT 720
AGACAAGGAA TTTCTCCAGG ACTTGAAGGA GCTCAAGGTG CTAGTGGCTG ACAAGGACCT 780
15 TCTGGACCTG CACAAGAGCC TGGTGTGCAC TGCTCTCCGG GGAAAGCTGG GCGTCTTCTC 840
TGAGATGGAA GCCAACTTCA AGAACCTGTC CCGGGGGCTG GTGAACGTGG CCGCCAAGCT 900
20 GACCCACAAT AAAGATGTCA GAGACCTGTT TGTGGACCTC GTGGAGAAGT TTGTGGAACC 960
CTGCCGCTCC GACCACTGGC CACTCAGCGA CGTGCGGTTC TTCTGAATC AGTATTCAGC 1020
25 GTCTGTCCAC TCCCTCGATG GCTTCCGACA CCAGGCCTCT GGGACCGCTA CATGGGCACC 1080
CTCCGCGGCT GCCTCCTGCG CCTGTATCAT GACTGAGGTG CCTCCCAACG CTCCGCCAC 1140
GCTGACAATA AAGTTGCTCT GAGTTTGAG ACTGGTCTC GCTCCGGGGA GCAAGTGGGG 1200
30 GGCCTGCAGA TGTCCCTGTG TCTGTCTCTG AGCACCTGGT GTCCGTGTAC AAGGATGGAT 1260
GTGTNCNGTG GCTCCTTGGG AACTGAGACA TATCTCAGGG AATGGTGTCT GTGCTCAGCC 1320
35 CATCCACCAG AAGA 1334

40 (2) INFORMATION FOR SEQ ID NO: 216:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1511 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 216:

50 GTGGCGGGGA TGCTGCGAGG GGTCTCCTG CCCAGGCGG GCCGGCTGCC TACCCTCCAG 60
ACTGTCCGCT ATGGCTCCAA GGCTGTTACC CGCCACCGTC GTGTGATGCA CTTCAGCGG 120
CAGAAGCTGA TGGCTGTGAC TGAATATATC CCCCCGAAAC CAGCCATCCA CCCATCATGC 180
55 CTGCCATCTC CTCCAGCCC CCCACAGGAG GAGATAGGCC TCATCAGGCT TCTCCGCCGG 240
GAGATAGCAG CAGTTTCCA GGACAACCGA ATGATAGCCG TCTGCCAGAA TGTGGCTCTG 300
60 AGTGCAGAGG ACAAGCTTCT TATGCGACAC CAGCTGCGGA AACACAAGAT CCTGATGAAG 360

	RTCTTCCCCA ACCAGGTCCT GAAGCCCTTC CTGGAGGATT CCAAGTACCA AAATCTGCTG	420
	CCCCTTTTCG TGGGGCACAA CATGCTGCTG GTCAGTGAAG AGCCCAAGGT CAAGGAGATG	480
5	GTACGGATCT TAAGGACTGT GCCATTCCCTG CCGCTGCTAG GTGGCTGCAT TGATGACACC	540
	ATCCTCAGCA GGCAGGGCTT TATCAACTAC TCCAAGCTCC CCAGCCTGCC CCTGGTGCAG	600
10	GGGAGAGCTT TAGGAGGCCT CACCTGCCTC ACAGCCCAGA CCCACTCCCT GCTCCAGCAC	660
	CAGCCCCCTC AGCTGACCAC CCTGTTGGAC CAGTACATCA GAGAGCAACG CGAGAAGGAT	720
	TCTGTCATGT CGGCCAATGG GAAGCCAGAT CCTGACACTG TTCCGGACTC GTAGCCAGCC	780
15	TGTTTAGCCA GCCCTGCGCA TAAATACACT CTGCGTTATT GGCTGTGCTC TCCTCAATGG	840
	GACATGTGGA AGAACTTGGG GTCGGGGAGT GTGTTTGTC CTTGGTTTTC ACTAGTAATG	900
20	ATATTGTCAG GTATAGGGCC ACTTGGAGAT GCAGAGGATT CCATTTCAGA TGTCAGTCAC	960
	CGGCTTCGTC CTTAGTTTTC CCAACTTGGG ACGTGATAGG AGCAAAGTCT CTCCATTCTC	1020
	CAGTCCAAG GCAGAGATCC TGAAAAGATA GGGCTATTGT CCCCTGCCTC CTTGGTCACT	1080
25	GCCTCTTGCT GCACGGGCTC CTGAGCCACC CCCTTGGGGC ACAACCTGCC ACTGCCACAG	1140
	TAGCTCAACC AAGCAGTTGT GCTGAGAATG GCACCTGGTG AGAGCCTGCT GTGTGCCAGG	1200
30	CTTGTGCTG AGTGCTGTAC ATGTATTAGT TCCTTTACTG CTGACCACAT TGTACCCATT	1260
	TCACAGAGAA GGAGCAGAGA AATTAAGTGG CTTGCTCAAG GTCATGCAGT TAGTAAGTGG	1320
	CAGAACAGGG ACTTGAACCA AGCCCTCTGC TCTGAAGACC GCGTCCTGAA TTTCTTCACT	1380
35	AGAGCTTCCT CATCAGGTTA CCCAGAAGTG GGTCCCATCC ACCATCCAGG TGTGCTTGGA	1440
	TGTTAGTTCT CCACCTCGA GGTGTACGCT GTGAAAAGTT TGGGAGCACT GCTTTATAAT	1500
40	AAAATGAAAT A	1511

(2) INFORMATION FOR SEQ ID NO: 217:

45

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 642 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 217:

55	AGGCCTTACT TTTCTCCCA CAAAGGAGTC GCAGCCACGC TAGCTCTGAC TTGCCACTGT	60
	GACAAAGTTC ACGTAGCAGG TCTAGGCAAA GACTGGGCAA TTGAGCAGAG GAGACGGACC	120
	TGTGAGTCTG ACCRYGAGSC GGRCCCTTTC ACCTTGGCTG GGCTGGTCTT GGTCTTAGG	180
60	TTTGTGTCAGG TTGTCCTTGT TTGGATCCCT CAACTAGGTG ATAAGCACTG GAGGGGGATG	240

5 ACCCGCCTTG GACGTGTTTC TTAACTCA TCCATATAAT AGGGCCGTGG GATGGTTGTA 300
GAGGTAAAGC AGGATGATGG TGTTTTAAGA CCAGAGCTTG GGACCAGGGC TCCTACACCT 360
AATTTTCTCT CCTGGTAGCT GAACAAAGGT CTAAATTAGC TTAACAAAAG AACAGGCTGC 420
CGTCAGCCAG AGTTCTGAAG GCCATGCTTT CAGTTTCCCT TGTGACAAT TGCTCTCCAG 480
10 TTCCTATGAA AGCACAGAGC CTTAGGGGGC CTGGCCACAG AACACAACCA TCTTAGGCCT 540
GAGCTGTGAA CAGCAGGGGG TTGTGTGTCT GTTCTGTTTC TCTGCTTGCC GAACTTTCTC 600
AATAAACCTT ATTCTTATT TTATATTAC GTNGGTGCTG GG 642
15

(2) INFORMATION FOR SEQ ID NO: 218:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1241 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

25

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 218:

30 GGTCCTACTG TTCCATTITA TGCTAATAGA TTCCATTCTA GGGCCAGCC GTCTCTTGAC 60
TGATGGTGTT CCCTTTAACC CTTGGCATGT ATAATAGAAT TTTGGTGAAT GAAAGAACCC 120
AATAGGCCA GATAGTCCCC CCAGGCCCTG ATATCCATAA AAGGCTTGGG AATGCATTAT 180
35 GTAATTGTCC TTAGTCTTTT TGTGTTTTA GAAAAAAAA ACAAGATGGG CTCAGATGGA 240
TGCTACGTA AAAATGGTTC CTAGCTGTGT ACTCATAACT TTTCTTTGAA TTGAGTAGTG 300
AAAGGAAGGA GGAGGAAAGG AAATTAATG TCCTTCTAGT ATTCTCTGGA CTCAAGTCTG 360
40 ACATATGRGA TAATAACCTA TATGAAATG CCAAGAATG TATCTGAAAC AAGRGAACAG 420
TTTGACACAT TTATCATGCC TTCATATTAC ATATTAACCTG AAACCAATTA ATAAACATAT 480
45 GAAATATCCA TTGCACAAGG CAAAGGCACC TAAACCTTTT GTTCTTTTCT CTACATAGCA 540
GAAATTGATT TTTTTTTTAT TTTTTTAGGG GAACCTATAT AATTATGACC CAGTGATGTC 600
TTTGGTGAC TTAAGCTTAT GAATTCAGGT TACAATTGAG TTGATTCTAG ATGGTTACTA 660
50 CCTTGAAAAG GATGTTGGTG CCTTATGTGA CACGAGCCAG AGCCTGCTGG GAATAAACAA 720
AGCAGATTCA TGCCAACACC AACTCGTAGC TTTAGTGGCA GATGGGAGTG GTCACAGACT 780
55 CCCAAAATGT GGGGCTTTGG ATTTCCACAC CATCCACGT GTGTGTCATC TTCTCTTTTC 840
ACACTCTTGA TGATAATTG AAAATGRTGA AATCACCTCT GAATTGCTCT ATAGCATGAG 900
CACATTCTTA TGACAACATA ACAAATAGTT CATAATGTGA ATATTAGAAA CTGTTACAGC 960
60

CTGCAGTTAC CATAATTTTC CATGTTTGTG GAATTGATAT TGAAATAGCA GGGCTAAGGA 1020
ATTACTGGCA AGTTTITAGCC TGIGGGTAAT ACCTTAGGGT TATTTAAATA TTTGTAATTT 1080
5 TATTTAAATG TTCATGAATG TTGAAAGGA ACAAATTAT CAGGGATGGC TCTTTGCCAT 1140
GGGTCTTATT TTCACCTCT TTTCTGTAAG AAAAAAGAAC AATGCTTAA TGTATTTTAA 1200
AAGTTTITGG TATAGTTTCT AATCCAAT TTAATAAAG T 1241
10

(2) INFORMATION FOR SEQ ID NO: 219:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1080 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

20

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 219:

25 TGTATTATGTG ACCTAAAACA TACACACATG CACACACACA TACATATCCA TTCATTCAAT 60
CATTCAGTG GTGTTTCCAG TGTCTGTGTG TCACTGTTTA TGCAGTTTCC ATTTCACAGT 120
GAATTATGAG TGGAGGGCAA CTTTCTAAC CAGATTGTCT TTTCAGAACA AAGACCKGGG 180
30 RATAGAGGAA GAGTTTGGAA AGAGGGAGAG GCAAGGAAAG AGAGCTTTAA ATTGAAAGGT 240
TAATTTCTTA AGAGGAACCT GGGCTGAATG ACTACAGTGT TATACCTCC AATCTTTGCA 300
GGTGGGCATG GAACACTGCT TGTATCACTC TGTGCACGGT ATAAATCCAT ATATCCACAA 360
35 AAACACACAT CCATCCATCA ACATATACAT GGTTTGGGAT GAGCAGGTCA ATAGTTTGA 420
GAGGGAGTTT GTTCTTTT TTTTCTCAT TACTCTTAA ATGTTGTCA GTTATCAAAC 480
40 AAACAAACAG AAAAATTGTT TGGAAAAAC CTTCATACG CCTTTCTAT CMAGTGCTTT 540
AAAATATAGA CTAAATACAC ACATCCTGCC AGTTTTTCT TACAGTGACA GTATCCTTAC 600
CTGCCATTAA ATATTAGCCT CGTATTTTC TCACGTATAT TTACCTGTGA CTTGTATTG 660
45 TTATTTAAAC AGGAAAAAA ACATTCAAAA AAAGAAAAAT TAACTGTAGC GCTTCATTAT 720
ACTATTATAT TATTATTATT ATTGTGACAT TTTGGAATAC TGTGAAGTTT TATCTCTTGC 780
50 ATATACTTTA TACGGAAGTA TTACGCCTTA AAAATACGAA AATAAATTTT ACAAGGTTTC 840
TGTTTTGTGT GGAAGAGTAA TTGATGTTGC TAAGAATGAT GTTGTGTTTT TTGGGGTTTT 900
TGTGTGTTTT TTTTAAATG TTACCAGCAC TTTTGTGTA AGTTTCACTT TCCGAGGTAT 960
55 TGTACAAGTT CACACTGTTT GTGAAGTTT AATATGAAG AATAATTAAA AAAAAAAAAA 1020
AAACCNCGGG GGGGGCCCG TCCCATGCGN CCAAGGGGG CGGTACGGG GTCACGGCCG 1080

60

(2) INFORMATION FOR SEQ ID NO: 220:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1258 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 220:

TGAATTGAGG GCTTAAAGAT AAACATATGG GRTTGGAGTT GTGTGTCCAT AGGGTTTCAC	60
15 TGCCTATTITG ATTTGAGTTT ATCCCTATTA ATTTTITACA GTGAAATTTT ATTAAAGTAT	120
AATGTACATA TATTTTCAGT GGATTTTGCT CTGAAGGTTT TCCAGTGGTC TGAATACGAG	180
20 ATAGTGGGGC TTCAGCTGTG GGATATTGCA GGGCAGGAGC GCTTCACCTC TATGACACGA	240
TTGTATTATC GGGATGCCTC TGCTGTGTIT ATTATGTTTG ACGTTACCAA TGCCACTACC	300
TTCAGCAACA GCCAGAGGTG GAAACAGGAC CTAGACAGCA AGCTCACACT ACCCAATGGA	360
25 GAGCCGGTGC CCTGCCTGCT CTTGGCCAAC AAGTGTGATC TGTCCCTTGT GGCAGTGAGC	420
CGGGASCAGA TTGACCGGTT CAGTAAAGAG AACGGTTTCA CAGGTTGGAC AGAAACATCA	480
30 GTCAAGGAGA AAAAAATAT TAATGAGGCT ATGAGAGTCC TCATTGAAAA GATGATGAGA	540
AATTCACAG AAGATATCAT GTCTTTGTCC ACCCAAGGGG ACTACATCAA TCTACAAACC	600
AAGTCTCCA GCTGGTCTG CTGCTAGTAG TGTITGGYTT ATTTTCCATC CCAGTTCTGG	660
35 GAGGTCTTTT AAGTCTCTC CCTTTGGTTG CCCACCTGAC MATTTTATTA AGTACATTTG	720
AATGTCTTCC TGAATCTGT CAGTAAGGA GGGCCATTGT CACTTAGAAA AGACACCTGG	780
40 AACCCAKGTG CATTTCTGCA TCTCTGGAT TAGCCTTTSA CATGTGCTG RCTCACATTA	840
GTGCCAGTTA GTGCCTTCGG TGTAAGATCT TCTCATCAGC CCTCAATTTG TGATCCGGAA	900
TTTGTGAGA AGGATKAGAA ATCAGCACCT GCGTTTITAGA GATCATAATT CTCACCTACT	960
45 TCTGAGCTTA TTTTTCATT TGATATTCAT TGATATCATG ACTTCCAATT GAGAGGAAAA	1020
TGAGATCAAA TGTCAATTCC CAAATTTCTT GTAGGCCGTT GTTTCAGATT CTTTCTGTCT	1080
50 TGGAAATGTAA ACATCTGATT CTGGAATGCA GAAGGAGGGG TCTGGGCATC TGTGGATTTT	1140
TGGCTACTAG AAGTGTCCCA GAAGTCACTG TATTTTITGAA ACTTCTAAGC TCATAATTAA	1200
GTTTCTCTTG TCTTGGGCAT CAAGANTAGT TCCAATTTT TGGGCCGGGG CAGGGTGG	1258

55

(2) INFORMATION FOR SEQ ID NO: 221:

60

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1693 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 221:

	CACAATATAT GAAATAGTAC CCTCTAAAAA AGAGAAAAAA AAAATCAGGC GGTCAAACTT	60
10	AGAGCAACAT TGTCTTATTA AAGCATAGTT TATTTCACTA GAAAAAATTT AATATCAAGG	120
	ACTATTACAT ACTTCATTAC TAGGAAGTTC TTTTAAAAAT GACACTTAAA ACAATCACTG	180
	AAAACCTGAT CCACATCACA CCCTGTTTAT TTTCCCTTAAA CATCTTGGA GCCTAAGCTT	240
15	CTGAGAATCA TGTGGCAAGT GTGATGGGCA GTAAAATACC AGAGAAGATG TTTAGTAGCA	300
	ATTAAAGGCT GTTTGCACCT TTAAGGACCA GCTGGGCTGT AGTGATTCCT GGGGCCAGAG	360
20	TGGCATTATG TTTTACAAA ATAATGACAT ATGTCACATG TTTGCATGTT TGTTTGCTTG	420
	TTGAATTTTT GAACAGCCAG TTGACCAATC ATAGAAAGTA TTAATTCTCT TCATATGGTT	480
	TTTGGTTTAC TGGCTTAAGA GGTTCCTCAG AATATCTATG GCCACAGCAG CATACCAGTT	540
25	TCCATCCTAA TAGGAATGAA ATTAATTTTG TATCTACTGA TAACAGAATC TGGGTCACAT	600
	GAAAAAAAT CATTTTATCC GTCTTTTAAG TATATGTTTA AAATAATAAT TTATGTGTCT	660
30	GCATATTGCA GAACAGCTCT GAGAGCAACA GTTCCCATT AACTCTTTCT GACCAATAGT	720
	GCTGGCACCG TTGCTTCCTC TTTGGAAGA GGAAAGGGTG TGTGAACATG GCTAACAATC	780
	TTCAAATACC CAAATTTGTA TAGCATAAAT AAAGTATTTA TTTTATGCCT CAGTATATTA	840
35	TTATTTAATT TTTTAGGTAA TGCCTATCTC TTGGTCTATT AAGGAAAGAA GCAATCAGTA	900
	GAGAATTCAG GATAGTTTGT TTTAAATCTT TGCAGATTAC ATGTTTTTAC AGTGGCCTGC	960
40	TATTGAGGAA AGGTATTCTT CYATACAAC TGTTTTAACC TTTGAGAACA TTGACAGAAA	1020
	TTATGCAATG GTTTGTTGAG ATACGGACTT GATGGTGCTG TTTAATCAGT TTGCTTCCAA	1080
	AGTGGCCTAC TCAAGAGGCC CTAAGACTGG TAGAAATTAA AAGGATTTCA AAAACTTTCT	1140
45	ATTCCTTTCT TAAACCTACC AGCAAAC TAGTGTGATA GCAATGAATG GTATGATGAA	1200
	GAAAGTTTGA CCAAATTTGT TTTTGTGTG TTGTTGTTGT TTTGAATTTG AAATCATTCT	1260
50	TATTCCTTTT AAGAATGTTT ATGTATGAGT GTGAAGATGC TAGCGAACCT ATGCTCAGAT	1320
	ATTCATCGTA AGTCTCCCTT CACCTGTTAC AGAGTTTCAG ATCGGTCACT GATAGTATGT	1380
	ATTTCTTTAG TAAGAATGTG TTAAAATTAC AATGATCTTT TAAAAAGATG ATGCAGTTCT	1440
55	GTATTTATTG TGCTGTGTCT GGTCTAAGT GGAGCCAAIT AAACAAGTTT CATATGTATT	1500
	TTTCCAGTGT TGAATCTCAC ACACTGTA CTGAAATTTT CCTTCCATCC TGAATAACGA	1560
60	ATAGAAGAGG CCATATATAT TGCCTCCTTA TCCTTGAGAT TTCACTACCT TTATGTTAAA	1620

AGTTGTGTAT AATTGTTAAA ATCTGTGAAA GAATAAAAAG TGGATTTAAA TTAAAAAAA 1680
AAAAA AAA 1693

5

(2) INFORMATION FOR SEQ ID NO: 222:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1196 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

15

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 222:

ACGCGTGGGT CGACCCACGC GTCCGCGACN TGGCGTGGTG GGAAGGGAG AAGGATTGT 60
AAACCCCGGA GCGAGGTTCCT GCTTACCGA GGCCGCTGCT GTGCGGAGAC CCCC GG GTGA 120
AGCCACGTC ATCATGTCTG ACCAGGAGGC AAAACCTCA ACTGAGGACT TGGGGGATAA 180
GAAGGAAGGT GAATATATTA AACTCAAAGT CATGGACAG GATAGCAGTG AGATTCACTT 240
CAAAGTGAAA ATGACAACAC ATCTCAAGAA ACTCAAAGAA TCATACTGTC AAAGACAGGG 300
TGTTCCAATG AATCACTCA GGTTCCTCTT TGAGGGTCAG AGAATTGCTG ATAATCATA 360
TCCAAAAGAA CTGGAATGG AGGAAGAAGA TGTGATTGAA GTTTATCAGG AACAAACGGG 420
GGGTCAITCA ACAGTTTAGA TATCTTTTTT ATTTTTTTTC TTTTCCCTCA ATCCTTTTTT 480
ATTTTTAAAA ATAGTTCCTT TGTAATGTGG GTTCAAAC GGAATTGAAA ACTGGCACCC 540
CATCTCTTTG AACATCTGG TAATTTGAAT TCTAGTGCTC ATTATTCATT ATTGTTTGT 600
TTCATTGTGC TGATTTTGG TGATCAAGCC TCAGTCCCT TCATATTACC CTCTCTTTT 660
TAAAAATTAC GTGTGCACAG AGAGTCACC TTTTCAGGA CATTGCATT TCAGGCTTGT 720
GGTGATAAAT AAGATCGACC AATGCAAGTG TTCATAATGA CTTTCCAATT GGCCCTGATG 780
TTCTAGCATG TGATTACTTC ACTCCTGGAC TGTGACTTTC AGTGGGAGAT GGAAGTTTTT 840
CAGAGAACTG AACTGTGGAA AAATGACCTT TCCTTAACTT GAAGCTACTT TTAAATTTG 900
AGGGTCTGGA CAAAAGAAG AGGAATATCA GGTGAAGTC AAGATGACAG ATAAGGTGAG 960
AGTAATGACT AACTCCAAAG ATGGCTTCAC TGAAGAAAAG GCATTTTAAG ATTTTTTAAA 1020
AATCTTGTC GAAGATCCA GAAAAGTTCT AATTTTCATT AGCAATTAAT AAAGCTATAC 1080
ATGCAGAAAT GAATACAACA GAACACTGCT CTTTTGATT TTATTTGTAC TTTTGGCCT 1140
GGGATATGG TTTTAAATGG ACATTGTCTG TACCAGCTTC ATTAAATAA ACAATA 1196

60

(2) INFORMATION FOR SEQ ID NO: 223:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 1791 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 223:

TCAGGGAGGT GGCAGGAAAG GCTTGAACA GCTGCCGAG TGACGGAGCG GCGGCCCGC 60
CCGGTTGCGC TGGAGTCTGA AGCTTCCAGG TAGCGGCCCG CAGAGCCTGA CCCAGGCTCT 120
15 GGACATCCTG AGCCCCAAGTC CCCCACTC AGTGCACTGA TGAGTGGGA AGTGAAGTG 180
ACAGGGCAGA ACCAGGAGCA ATTTCTGCTC CTAGCCAAGT CGGCCAAGGG GGCAGCGCTG 240
20 GCCACACTCA TCCATCAGGT GCTGGAGGCC CTTGGTGTCT ACGTGTMTGG AGAACTGCTG 300
GACATGCCCC ATGTTAGAGA GCTGGCTGAG AGTGACTTTG CCTCTACCTT CCGGCTGCTC 360
ACAGTGTTCG CTTATGGGAC ATACGCTGAC TACTTAGCTG AAGCCCGAA TCTTCCTCCA 420
25 CTAACAGAGG CTCAGAAGAA TAAGCTTCGA CACCTCTCAG TTGTCACCCT GGCTGCTAAA 480
GTAAAGTGTA TCCCATATGC AGTGTGCTG GAGGTCTTGC CTGCGTAAT GTGCGGCAGC 540
30 TGGAAGACCT TGTGATTGAG GCTGTGTATG CTGACGTGCT TCGTGGCTCC CTGGACCAGC 600
GCAACCAGCG GCTCGAGGTT GACTACAGCA TCGGGCGGGA CATCCAGCGC CAGGACCTCA 660
GTGCCATTGC CCGAACCTTG CAGGAATGGT GTGTGGGCTG TRAGGTGCTG CTGTCAGGCA 720
35 TTGAGGAGCA GGTGAGCCGT GCCAACCAAC ACAAGGAGCA GCAGCTGGGC CTGAAGCAGC 780
AGATTGAGAG TGAGGTTGCC AACCTTAAAA AAACCATTA AGTTACGACG GCAGCAGCAG 840
40 CCGCAGCCAC ATCTCAGGAC CCTGAGCAAC ACCTGACTGA GCTGAGGAA CCAGCTCCTG 900
GCACCAACCA GCGCCASCCA GCAAGAAAGC CTCAAAGGGC AAGGGGCTCC GAGGGAGCGC 960
CAAGATTTGG TCCAAGTCGA ATGTAAAGRA CTGTCGTTTC CTCCTGGGG ATGTGGGGTC 1020
45 CCAGCTGCCT GCCTGCCTCT TAGGAGTCCT CAGAGAGCCT TCTGTGCCCC TGGCCAGCTG 1080
ATAATCCTAG GTTCATGACC CTTACCTCC CTAACCCCA AACATAGATC ACACCTTCTC 1140
50 TAGGGAGGAG KCAAATGTAG GTCATGTTT TGTGGTACT TTCTGTTTTT TGTGACTTCA 1200
TGTGTTCCAT TGCTCCCCGC TGCCATGCTC TCTCCCTTGT TTCTTAAGA GCTCAGCATC 1260
TGTCCCTGTT CATTACATGT CATTGAGTAG GTGGGTAGCC CTGATGGGG TCGCTCTGTC 1320
55 TGGAGCATAA CCCACAGCG TTTTCTGTC CCCCCATCC CTGCATGCCT GATCCCCAGT 1380
TCCTATACCC TACCCCTGAC CTATTGAGCA GCCTCTGAAG AGCCATAGGG CCCCCACCTT 1440
60 TACTCACACC CTGAGAATTC TGGGAGCCAG TCTGCCATGC CAGGAGTCAC TGGACATGTT 1500

15
5
10

CATCCTAGAA TCCTGTCACA CTACAGTCAT TTCCTTTCCT CTCTCTGGCC CTGGGTCCT 1560
GGGAATGCTG CTGCTTCAAC CCCAGAGCCT AAGAATGGCA GCCGTTCCTT AACATGTTGA 1620
GAGATGATTC TTCTTGGCC CTGGCCATCT CGGGAAGCTT GATGGCAATC CTGGAAGGGT 1680
TTAATCTCCT TTGTGAGTT TGGTGGGAA GGAAGGGTA TATAGATTGT ATTAAAAAA 1740
AAAAGGTATA TATGCATATA TCTATATATA ATATGACGCA GAAATAAATC T 1791

15 (2) INFORMATION FOR SEQ ID NO: 224:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 2517 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 224:

25 ACACTAGTGG ATCCAAGAA TTCGGCACAG CGGCACAGCA TTGTGAGCT TTTCTGTGTG 60
TGTGGGGCCC TCAAGCGAGC TCGACTGGTC CATCCTGGGG TAGCGASGTG GTGTTTGTGA 120
AAAAGGACGA TGCCATCACC GCATAYAAGA AGTACAACAA CCGGTGTCTG GACGGGCAGC 180
30 CGATGAAGTG CAACCTTCAC ATGAATGGGA ATGTTATCAC CTCAGACCAG CCCATCCTGC 240
TGC GGCTGAG TGACAGCCA TCAATGAAA AGGAGAGCGA GCTGCCTCGC AGGGTGAAC 300
35 CTGCCTCCTC CTCCAACCCC CCGCCGAAG TGGACCCTGA CACCATCCTG AAGGCACTCT 360
TCAAGTCTC AGGGGCTCT KTGACCAGC AGCCACAGA WTTCAAAATC AAGCTTTGAG 420
CAGGGGAGTR AGGCAGCCAG AAGTGGGGC AGAGGAGGGT GGCTCTGTTT CCCCAAGGCA 480
40 AAGCTTATGA CCAATGGGCC ATCGGACTGG AGACCCCTGA TTGTGGGAAG GGTGCCAGG 540
GATAAAGAGC TTCTCACTG GATGGGACCC GCCTTCTGT GTTGTGTTCT GCCCTGTGCT 600
45 CTCTCTCTA CGTTAAGTT TCCTGTAGTA TGTTCCTCA TCTCATCGCC AAGGTAGGCT 660
TGTGTTTTTM AGTGTGTGCC TCCCCGAGCC TCAGCCCCAA GCTGATTCTT TATCTGGAAA 720
TGGTACACTG AATTCTCTGG GTGGCTTCT TGTGGCCCA TGGGATGCAG CGTGGGGGCT 780
50 GTCTGAAGGA CCCTGCTTTT TCCAGGGGCC GAGGGGCTGC CTTTCCTTTG TGTGTATTAA 840
GCTTTTCAA CAATGGAGGG GATGGAGAGC CCTGGTGTCC TGACGGGAGC CAGGTCCGCC 900
55 TGAGAGCTGT GCCGCTCCTC TGTCTGTCA GTGGAGGTGC CTGGGTGGGG AGCAGGTCTC 960
AGGCCTCTTG TCCTCTCCCC AGTGGCTCCA GGCCTCACTA GTGGCAAGGG CAGGATGAGG 1020
CTGCACCGCT GGAAGAGTC TATCTAAGCT CTTGGCTTGG AGTCCCGTGT CGTCTCCRCC 1080
60

CAGAGGAAGT TCTCCAGAGT TCACCTTTCC CTTTTCTTG AGTTGTGCTG AATGCCCCAC 1140
CCCAGCTCTC TTTCCTTCT GGTGTCTTT GCTGGGAGGG GGCTGTGTG TGAGCCCTCC 1200
5 CGGTCTCAC CTCGCCTGGC ACTTAACCAC ACCCTGGTTT TGTGTAGCCG CCAGCTCTCT 1260
TCTGGTTGGG CCTTTGAAAG GCTCAGCTC CCATTGTGCA GTGCTTGGGT TTGGAGCTTA 1320
10 TTTGAATGGA AGAGGTCAGT TTGTTCTGG CTCTCCATTT CTGGCCTCAG TTGTCTACAG 1380
GACAGTGGTC AGGGATGCCT GGAGGCATAT ATCCAGCTGC CACCAAGGGG CACTGTTTGT 1440
TCCCACTTAT GTGAGTGACC CCATCCATCC ATGACCAGAG GATTATTTTC CTGCCTTGGC 1500
15 AGAGGAGGAG GAGTCAAGGG AGCAGGGCAG CTCTACCAGG CAAGGTGTTT CCCAGCATA 1560
GGCGCAGACA GTTGGGACGA AACTTCAGAG CCCAGGCAGT CCCTGAATGA CCAGGCCAGT 1620
GTTGTCACTG AGTGGTCCCC TGCTGGTTGG GAGTGAAGAG AATCCAGGCT GGCAGAGCTG 1680
20 GAGCCAGTTG GGGAGCACGG TTCTGGGAGC TCTGCAAAAT CAGTAGCAAG TGCTGGAAAA 1740
GGCACATGCC GAAGATACTC AAGAGCTCCC AAGATTTGCT TGAGGCTAGC CCAGTAAAA 1800
25 AAACAGAGA CTCATGTTTC CAGGGTTCAG TCTGTCAAGC AGGAAGGACC CAGGATTGA 1860
ACCCAGCTTC AGTGTGCAGG CTCTGAGGCT GCCCAGGACG GGAAAGTCCA AGGAAGGGGC 1920
CTGGTGGTGC TCCACTTGCA GTTCTTTAAA GAATGCTGCT TTTTATCTC CTAACCTTT 1980
30 CAAGTGGGTG CAGACTTCTC GTTAGCAGCT GGAAGACATT CCTCCACAC TTTTCCCTTC 2040
CTGGCCCAAG AGAGCATCCA GAAGGCAGTA GGACCTGGTT TTTCAGGTAC TGGGAGCCGG 2100
35 GGGTCACTG CTTGCACTGT GCTTAGGGTA GGGATGGTAA ATATCCTCCC TGCATGGCTT 2160
TATCCTCCCT CTCATCCCAA AGCAGGTATC TTCTGGTTGT CACAGAGTTT CATTGAGTCC 2220
AGCTGCAGCC ACGTGGCCAT CTGGAGCTGG TGCTATAGGT GACCATCTGG TACATTGAGG 2280
40 GGACCTGTTT GCCTCCTCCA CTCTATAAGC AGTCATCTTG GGAGACGGG AGGAGAAGGT 2340
GGTGGGCTAG TCCTGTGTCC TCCTCCACTT CCCATGCCTC TATGTTACCC ATCTGTGTCT 2400
45 CCTGTGCAGA AGGAGAGGAA GGGGCATTAA GAGATGAAG GTGATTATGT ATTACTTATC 2460
CATTCTGAA TAAACATTTG TTATTCCTAA AAAAAAAAAA AAAAACTCG AGGGGGG 2517

50

(2) INFORMATION FOR SEQ ID NO: 225:

(i) SEQUENCE CHARACTERISTICS:
55 (A) LENGTH: 2424 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 225:

	TTGTANCTAA TCGAGGATTG ATTCTAATGA CAGAGTCTTT CAACACTTTG CACATGATGT	60
5	ATCACGAAGC TACAGCTTGC CATGTGACTG GAGATTTAGT AGAACTTCTG TCAATATTTT	120
	TTTCGGTTTT GAAGTCTACA CGCCCTTATC TTCAGAGAAA AGATGTGAAA CAAGCATTAA	180
	TCCAGTGGCA GGAGCGAATT GAATTTGCC ATAACTGTT AACTCTTCTT AATTCTATA	240
10	GTCTCCAGA ACTTAGAAAT GCCTGTATAG ATGTCCTCAA GGAAGTTGTA CTTTGTAGTC	300
	CCCATGATTT TYTTCATACT CTGGTCCCTT TTCTACAACA CAACCATTTG ACTTACCATC	360
15	ACAGTAATAT ACCAATGTCT CTTGGACCTT ATTTCCCTTG TCRAGAAAAT ATCAAGCTAA	420
	TAGGAGGGAA AAGCAATATT CGGCCTCCGC GCCCTGAACT CAATATGTGC CTCTTGCCCA	480
	CAATGGTGG AACCAGTAAG GGCAAAGATG ACGTTTATGA TCGTATGCTG CTAGACTACT	540
20	TCTTTTCTTA TCATCAGTTC ATCCATCTAT TATGCCGAGT TGCAATCAAC TGTGAAAAAT	600
	TTACTGAAAC ATTAGTTAAG CTGAGTGTCC TAGTTGCCTA TGAAGGTTTG CCACTTCATC	660
25	TTGCACTGTT CCCCAAACCT TGGACTGAGC TATGCCAGAC TCAGTCTGCT ATGTCAAAAA	720
	ACTGCATCAA GCTTTTGTGT GAAGATCCTG TTTCGCAGA ATATATTAAA TGTATCCTAA	780
	TGGATGAAAG AACTTTTTTA AACAACAACA TTGTCTACAC GTTCATGACA CATTTCTTTC	840
30	TAAAGGTTC AAGTCAAGTG TTTTCTGAAG CAAACTGTGC CAATTTGATC AGCACTCTTA	900
	TTACAAACCT GATAAGCCAG TATCAGAACC TACAGTCTGA TTTCTCCAAC CGAGTTGAAA	960
35	TTTCCAAAGC AAGTGCTTCT TTAATGGGG ACCTGAGGGC ACTCGCTTTG CTCCTGTGAG	1020
	TACACACTCC CAAACAGTTA AACCAGCTC TAATTCCAAC TCTGCAAGAG CTTTAAAGCA	1080
	AATGCAGGAC TTGTCTGCAA CAGAGAACT CACTCCAAGA GCAAGAAGCC AAAGAAAGAA	1140
40	AAACTAAAGA TGATGAAGGA GCAACTCCCA TTAAGAGGCG GCGTGTAGC AGTGATGAGG	1200
	AGCACACTGT AGACAGCTGC ATCAGTGACA TGAAACAGA AACCAGGGAG GTCCTGACCC	1260
45	CAACGAGCAC TTCTGACAA T GAGACCAGAG ACTCCTCAAT TATTGATCCA GGAAGTGAAC	1320
	AAGATCTTCC TTCCCTGAA AATAGTTCTG TTAAGAATA CCGAATGGAA GTTCCATCTT	1380
	CGTTTTCAGA AGACATGTCA AATATCAGGT CACAGCATGC AGAAGAAGAG TCCAACAATG	1440
50	GTAGATATGA CGATTGTAAA GAATTTAAAG ACCTCCACTG TTCCAAGGAT TCTACCTTAG	1500
	CCGAGGAAGA ATCTGAGTTC CCTTCTACTT CTATCTCTGC AGTTCTGTCT GACTTAGCTG	1560
55	ACTTGAGAAG CTGTGATGGC CAAGCTTTGC CCTCCAGGA CCCTGAGGTT GCTTTATCTC	1620
	TCAGTTGTGG CCATTCCAGA GGAAGTCTTA GTCATATGCA GCAACATGAC ATTTTAGATA	1680
	CCCTGTGTAG GACCATTGAA TCTACAATCC ATGTCGTAC AAGGGATATC TGGCAAAGGA	1740
60	AACCAAGCTG CTTCTTGACA TTAGGTGTAG CATGTCTACT TTAAAGTCCC TCACCCCCAA	1800

CCCCCATGCT GTTTGTATAA GTTTTGCTTA TTTGTTTTTG TGCTTCAGTT TGTCCAGTGC 1860
 TCTCTGCTTG AATGGCAAGA TAGATTTATA GGCTTAATTC TTGGTCAGGC AGAACTCCAG 1920
 5 ATGAAAAAAA CTGTCATCTT CAGTATACTT CCTAAAGGGC AATCAGATAA TGGATATGTT 1980
 TTATGTAATT AAGAGTTCAC TTTAGTGGCT TTCATTTAAT ATGGCTGTCT GGAAGAACA 2040
 10 GGGTGCCTA GCCCTGTACA ATGTAATTTA AACTTACAGC ATTTTACTG TGTATGATAT 2100
 GGTGTCCTCT GTGCCAGTTT TGTACCTTAT AGAGGCAGAT TGCTCCGAT CGCTGTGGTT 2160
 CTTATTATCA AAATTAAGTT TACTTGATA CGGAACAACC ACAAGAAAT TGAATCTGTA 2220
 15 AAGAATCCTC TTTAGCTGTG GCCTGGCAGT ATATAAATGG TGCTTTATTT AACAGAATAC 2280
 CTGTGGAGGA AATAAAGCAC ACTTGATGTA AAAATAATTG TTTTATTTT ATTGACATGA 2340
 20 CTGATTGATT GCTATTCTGT GCACTNAATT AAAGTATTG TGATGACTTA AAAAAAAAAA 2400
 AAAAAAAAAA AAAAAAAAAA AAAA 2424

25

(2) INFORMATION FOR SEQ ID NO: 226:

(i) SEQUENCE CHARACTERISTICS:
 30 (A) LENGTH: 1080 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 226:

ATATAGGACG GATAATCTGT TTACATTCTG TTCTTCTCGA TGCACTCACA AGCGGGTAAC 60
 TAGGTGACAA GAAACAAAG ATCTTATTCA AAAGAGGTCT TACAGCAACC CAACGTCTCA 120
 40 TCTTCCATA GTAAAGATGA CGGCGCCTTG AGGTAAGCTA CAGGCAACAC CACTTCCGCG 180
 TTTCTCTTGC GCCCTGGTCC AAGATGGCGG ATGAAGCCAC GCGACGTGTT GTGTCTGAGA 240
 45 TCCCGGTGCT GAAGACTAAC GCCGACCCC GAGATCGTGA GTTGTGGGTG CAGCGACTGA 300
 AGGAGGAATA TCAGTCCCTT ATCCGGTATG TGGAGAACA CAAGATGCT GACAACGATT 360
 GGTCCGACT GGAGTCCAAC AAGGAAGGAA CTCGGTGGTT TGGAAAATGC TGGTATATCC 420
 50 ATGACCTCCT GAAATATGAG TTTGACATCG AGTTTGACAT TCCTATCACA TATCCTACTA 480
 CTGCCCCAGA AATTGCAGTT CCTGAGCTGG ATGGAAAGAC AGCAAAGATG TACAGGGGTG 540
 55 GCAAAATATG CCTGACGGAT CATTTCAAAC CTTTGTGGGC CAGGAATGTG CCCAAATTTG 600
 GACTAGCTCA TCTCATGGCT CTGGGGCTGG GTCCATGGCT GGCAGTGGAA ATCCCTGATC 660
 TGATTAGAA GGGCGTCATC CAACACAAAG AGAAATGCAA CCAATGAAGA ATCAAGCCAC 720
 60

	TGAGGCAGGG CAGAGGGACC TTGATAGGC TACGATACTA TTTTCCTGTG CATCACACTT	780
	AACTCATCTA ACTGCTTCCC CGGACACCCT CCACCTCTAG TTGTTACTAA GTAGCTGCAG	840
5	TAGGCATTGC TGGGGAAGAA ACAAAACACAC ACCAAACAGT ACTGCTACTT AGTTTCTAAG	900
	GCTGCACAGG GAAGGGAAAG ACTGGGCTTT GGACAATCTA GAGGTAATTT ATATCCGCCC	960
10	CCAGGTGGAG CAACATGCCA TTCTGGAGGC ACGGGGGTAA CTGAAAGTGA GTACATATAG	1020
	TCTTCTGGT TTCTGGAGAT AACCCATCAA TAAAAGCTGC TTCCTCTGGG TAAAAAAAAG	1080
15	(2) INFORMATION FOR SEQ ID NO: 227:	
	(i) SEQUENCE CHARACTERISTICS:	
20	(A) LENGTH: 1336 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 227:	
	TTGCATTAC AATTACTGGG AGGCAGGCAG GGCAGTTGC ATGCTGGGGG TGGCTGCATG	60
	GSCTGCCASC TCTCCTGGGT TTGAAGGATG CGGTACASCT GCTTCAGCTG AGCAACGATG	120
30	TTATCCTTGA TGTCTGGGGT TGAGATCTGC AGGCGGACAC TGCCACTATC AAAGGATCGT	180
	GTGAAATCAC CAGAAAACAT CTCGTAGATC ATCCGAGCCA CTA CTGGAAT GACCTGAACC	240
35	AAGATGAGTT TCCTTTCCAA TGGTTTCCCA TCTGGCCATT CTTCGCCAAA GCATAAGTAG	300
	ATCTCAAACG GTGGCTGCTT CTCTATCTGT CCTTCTGGT GGGCAATGAG ATCGCTAAGG	360
	AATGTTTCCA GACAAAATAG CTTGACCTTC TTTGTCTCT CAATCAGGTT GGGAGCAACA	420
40	AGTGATGGGG CACATGGCCC AGACCAGTAC ACCTTGCACT GGCACAGYCT GATGGCATAA	480
	ATGGCATGAC CGCTGACCTC CAGGATCAGT CCTCTGTCCA TGACGTCCAG CAGCTTGCTA	540
45	GTGAACAGCT TCTGCTTCTC ATTGGAATA TGCTCAGGAC CTGGGAATTT GACCTGCTCC	600
	AGNCTGACGG GACCAAAGAG CTCCTCCTGG TCAGGCATGG GACCCAGGTC CCCATAGAAG	660
	AGTCGGCAGC CCTGAGGGTT GCTCACGGTC ATGGTCCTGC CCGTACTCCT TCCCACGGTA	720
50	CTGAAACTTG ATGTCCAGGT CAGTCATTGG GAGAGAGCTG ATCCACAGTT CTGGAGAGCT	780
	ATAGAAGGRC TGTATAGGTG CCTGGGGWAC TTCCATCTCC AGGGGTTGAG TTTTGGGCCA	840
55	CACTGCCTCC GGSCTGCAGT TGCCACACT GCAATTGCCC ACACTGGCTG GCGCCATGGG	900
	AGAACCATTG ATGTTCAAGG AGGGGAAGGT GTCCTGGATG GGAACATGGT GCTGCCACTG	960
	ATCCAGCTCA TCTTCCTCAT CTCTTCATC CACATCATTA TCCTTCTCAT CCCAGGGAGC	1020
60	AGACCCTGTG GATCCTGGGT TAATGATCGA SCCCTGGGGC TGAGGGATGT CACACACTTG	1080

ATATATCTTC ACTGGGTCA TGGGCACCTC CCTTGGTGCC ATCCATACAT CCAGGTGAA 1140
 TTCTCTGCTC TTATGAGAG CACAGCGCAG CTGGGCCTTC CATTTAGCTG GGTGAGGGTC 1200
 5 ATCCACCCCT TCCTGGTACT TCCCTGTCTC TACAGCCCAG GCCTTAAAAA TGGTATTTTC 1260
 CTCTTCTGTG TGAGGGCTAT GCCGGGTGGC ATGPTTCCAG GGAATCTGGA AGCGTTTGA 1320
 10 GTCCCTGTGT AGCCAG 1336

15 (2) INFORMATION FOR SEQ ID NO: 228:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2043 base pairs

(B) TYPE: nucleic acid

20 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 228:

25 TCAGCTGGTC CCTTCCTTGT GTCTGGGGG ACCTGCTGGC GGCCTCTTCC TGGGAGCCAT 60
 GACCTCAGAC CCCACCCACA CTCCAGATCG AGACCCCTGC CTCCTCCCGG CAAATGTCCT 120
 CCCGCTGCCT TGCAGCTGC ACTTTGCACA TGCTACCCC CAGCACAGTC CCACTGGCCC 180
 30 CTCAMCTCCC CTCCCTGAG CTCCTTCCA AGGACTCCTG GTCAGTGCCT GCTGTGCART 240
 CAGAGGCCCA GGTCCAGCA GCCCGGSGGG AACGGGTGCT GCCTSTTCCT CCAGTTAGCT 300
 35 CCAGYTCAGG TCTGAGACC GTGYTGAGTA AAGTCTGAG CAMCGACCGT GCCCTCTGCC 360
 CAGGGCTGGG TCCTGAGCAG CTGGTTTTC TGAGGAAGG TTGGAGCAAG CAAAGTCCTT 420
 CTCTGCCCTC AGGGTCAGCT GCCCAGACTG GGGCGGATGC AGAGAGGCAG GTGGGCTGTG 480
 40 GCTGGACTGG TCCGAGCTG GCTTCCTTAC CAGAAAAGCC TCAGCCTTCC TCTGGAAGCA 540
 TCCCCCGTTC TGGCAAGGG GGAAGGGCTC CTTTAAGGGG TGTGCTTTC CAGTGGGGAG 600
 45 CAGTCTGGCC CTGCCCCCTA CTAAAGCCTC TGCTCTCAGC ACTTTCCCCC AAGTCCTTGT 660
 AACTTGCTTG AAGGTGGGTT CTGGCTGCCA GCCAGTCCCT GGACAAACTC TCCTGCCCTT 720
 TTTAAATTT ACTCATTTG TATAAACCCA GCAGGCTGGT GTTACTTAG CCCTGTAGCT 780
 50 TTTTTCATTT TTTCTTTCCG TCTTCTTCT TGAGTTCAGG GTTCAATATT GCCTCCTCGC 840
 CCTGGTGAGG GGAGGTGCTG CTTTCTGCC CCACCTGCCG GCTGGTTCCA GCAGCGCTGG 900
 55 NGCCAGCTG GGGGGCCGGG ATGGGGCTT CTCTCTCTGG GAGGGGTGCA GGTGCCCTCC 960
 CCAGGCTGGG AGGGTTCCTT CCCTAGCTCC CCACTGCCCC CCGCTGGTGA GAGTTGGGCT 1020
 60 TCTTGGTCTT GGAATCCCT GGCATTGGGA ACAGAGCATT TCCAGCATTT GTTGTGTGTG 1080

TTTTACTCAC CTAACCCCTTA GAAAATGAAT GTTAGAAGGT GCCTGCCGAG GCGGGACAGA 1140
 GTGTTTGCTC GCGCTGGAGA AGGCTCTGCT CAGCCCTGAG AGTCCCTTCC TGCCCCACCG 1200
 5 ATACTGGCAC TTTAAAAAGG AAGCTGACCG CACAGTGTCC AGACGAATTG GCCCCCAGAA 1260
 GATGGGGAGT TCTGTCTGC CCTTCTGTGT CTGCGTGACC TCACCAGCC TAGGAGGGAG 1320
 10 GTGCATTTCAG GGTAGATTG CCTCTCATT CAAAGTCTGG GCCTTTGGGY GGAAAACAGC 1380
 CAGCTTTGGC GCTGTTGGG AGACTCCTCC AGACCAGGAA CCCCAGAAG AGACAGAGCC 1440
 TGCCACATCC TCCACGCCA GGCCTGGGC CAGGGTGATT GGA CTGAGAA TTTGGCCACA 1500
 15 ACCAAATTGA TGCTGGCTGG AACCAGAGGC CAGAAAGCCT GCCTTGTCC CCATGTGGGA 1560
 GCCTGTCTT CAGCCCTCTT GTCCCTTGA GCTCAGTGAA TCCCACCAG GTGCCACAG 1620
 CTCTGGACT TCAAATCTA TATATTGAGA GAGTTGGAGA GTATATCAGA GATATTTTTG 1680
 20 GAAAGGAGTT GGTCTATGCA ATGTCAGTTT GGAATCTTCT TGAAAGTTTA ATGTTTTTAT 1740
 TAGGAGATTT AAAGAAAATA AAGGTCTACA ATATCTTTAG GTTTTTTTTT TTTCTGTIT 1800
 25 ACCGCACAAA CTGACCACAT GGCATGTCTA TCAGGATGGA GGGTGTCCAT GTTCTCCTCT 1860
 GTCTTTAGGG AGGTGATAAG GAGATGGSCG RAGGGGTGTT TTTTCTTTG ACTCCCTCC 1920
 TTTCTAACAG AATGTTGCCA CCACTGCTTG AGTGGGCTGT GTTTGTTCTT CTGTCCACG 1980
 30 TTCTGTGTGA GAAATAACA TTGTTAGGGG AACTCAGGCT AGTGTACAGG TCTTGTTTG 2040
 GGG 2043

35

(2) INFORMATION FOR SEQ ID NO: 229:

40

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 540 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 229:

TAAAAAGAAG CGGAGAAATC TGGGCGTCGC TCTAGAGATC GATGGGCTAG AGGAGAAGCT 60
 50 GTCCCACTGT CGGAGAGACC TGGAGGCCGT GAACTCCAGA CTCCACAGCC GGGAGCTGAG 120
 CCCAGAGGCC AGGAGGTCCC TGGAGAAGGA GAAAAACAGC CTAATGAACA AAGCCTCCAA 180
 CTACGAGAAG GAACTGAAGT TTCTTCGGCA AGAGAACCGG AAGAACATGC TGCTCTCTGT 240
 55 GGCCATCTTT ATCTCTCTGA CGCTCGTCTA TGCCTACTGG ACCATGTGAG CCTGGCACTT 300
 CCCACAACC AGCACAGGCT TCCACTTGGC CCCTTGGTCA GGATCAAGCA GGCACCTCAA 360
 60 GCCTCAATAG GACCAAGGTG CTGGGGTGTT CCCCTCCCAA CCTAGTGTTC AAGCATGGCT 420

TCCTGGCGGC CCAGGCCTTG CCTCCCTGGC CTGCTGGGGG GTTCCGGGTC TCCAGAAGGA 480
CATGGTGTCTG GTCCCTCCCT TAGCCCAAGG GAGAGGCAWT AAAGACACAA AGCTGGAAAT 540

5

(2) INFORMATION FOR SEQ ID NO: 230:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 448 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

15

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 230:

AATGTGTGAAA TATTAGAATA TTGTTACTAT TTGACCCAAC TCAAAATCTC CATGGGAAAA 60
TACCTGTGCGA TACCCACAGT ATTGTTGAAA ATAATCAGAT GCAGTATCAC AGCTGTGTCA 120
GACTCTAGTA CCAGTTGGGC AATCAAGGCA CAGCTAAAAA TTGAAAACAA AGATCTGGAC 180
25 AACAAAACAG CCAAAGGTGG GGGTCAAGAA GCTCTGACGT GTACCTAGCT GTAGAATGCT 240
ATGCACACGT GCCAGGTGTA GTGTGCATAT CCAGGAAAAA CTGCAGAGAG CCCAGTCTT 300
CAMCTCTGGT TGACCATGAG CTCTGTGTAA GCAGGAAGTG AAGGCTAAGG CAGATTTAAG 360
30 CTCTGAAAGC ATTCCACAAC ATACACACAA ATCGTGCAAA GCATTAAGGA AATCTTGTTA 420
CTGCTAAGTG TTGCTGACCC AGGAACAA 448

35

(2) INFORMATION FOR SEQ ID NO: 231:

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 407 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

45

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 231:

GTATGCTGCC CCAAACCAAT ATGTGTGGCT GCCTTTWACC TGACTTCTCC AACATGTAGC 60
50 CCCAAGAGGA GGCCTCTAGA CTRAGGGAGG GGCTGGTGAC CCAGGTGTGG TGGGGCTGCA 120
TGARACTACC AGAGAGACAG ACATTCTGGA ACTCACCCCTG GGGGATCCAG TGGATCTGCC 180
TATGTTCTGG TCCACCCACG ACCTGTGAGA TGTTCCTCAT GAGGATGCAC TTGTGCTTCT 240
55 GCAAGTATTG CTGCAGCTTC ATAGTGACTC CCACCAGCAC CAGCAATACA GYTAGCTACC 300
TGTTGGCCTTG GATCTCAGCC AGCATGGCTG GGAGAGGGAG CARCTGGGCA TGTACCCTAA 360
60 ATGCTGTTAC CAGGAAGGA CTCCAGAGT GAAGACAAGT AGGGACT 407

5 (2) INFORMATION FOR SEQ ID NO: 232:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 830 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 232:

15 GTATTGATT TCAGGCTGCT AAATGGGCTC ATTTAGCATT CATTCCCTGA TGTAGACATT 60
AAAAAAAA CTGAATAGCA TTCTTTCCAG GNTAACTAAT AAAGCAGACA TGCTAAGCCT 120
ATAAATACAT CAGCACTGCA GCACACGTTT AAGGTTGCCA CGGACAAGGA TCACACAATA 180
20 GAGAACACTG TAGTTCGGTC TGCTCACAAG ACCCAGAACA TTGATCAGTT TTTGTTGTTG 240
GTTTATTATT TTCTGTAA AAAATTGTGA AAAGTTTGT TTAGCTAGAT GATATTTTAA 300
25 TAGCTGCGAG TGCTTTGGAA CTATAAAGAT GTCACTACTT AACACACATA CCTTATGTTT 360
TGTTTTGTTT TGTTTTACAC TCAGTATAAA TCAGGAGAAG TTAGCCAACC ATCTAGCATT 420
TAGAATCCTC TTTTTTATTG TCTTCTAAGG ATATGGATGT TCCCATAACA GCAACAAAAC 480
30 AGCAACAAA ACATTTCATA AATATCACTT GATAGACTGT AAGCACCTGC TTAACTTTGT 540
GTNCCAAATA TTTAGTGTGT ATATATATAT ATATATACAC ACACACACAC ATATATATTC 600
35 AACAAATAAA GCAAATATA ACATGCATTT CACATTTTGT CTTTCCCTGT TACGATTTTA 660
ATAGCAGAAC TGTATGACAA GTTAGGTGA TCCTAGCATA TGTTAAATTC AAATTAATGT 720
AAAACAGATT AACAACAACA AAGAACTGT CTATTTGAGT GAAGTCATGC TTTCTATTAT 780
40 AATAACTTGG CTTCGGTTAT CCATCAAATG CACACTTATA CTGTTATCTG 830

45

(2) INFORMATION FOR SEQ ID NO: 233:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 932 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 233:

55 CCAGAAGAAA GACCAATCTA GAATATGGAA CTCTAATCAC TTCTAGTATT TCAACTTCCT 60
AGCAGAAATG AACTTGGCCC TAGACCTAGG GGATAAGCAA TGTCTTTTAT GTAGCCAATG 120
60 CTACGGAAAC AAAAGAGGTG AAAGAGACCC TTTTTTTATA CTTAATGTAC ATATATTGAC 180

TTTTTGAGCA AGAATGCCAG AAATAGCCTT CATTTCTACC CTGCAAATA ATCCAGATCT 240
GCTTTCTAAA ATGRANTCAG TTTCTAAAGT GAAACATGCA ATATTTATGC TCTGACTGAC 300
5 TCCTGAATTG GARGAGGAAG RACTTCTGTT TACAGAAAAC YGTATTGTTA TATATGTCAG 360
GCTGTGTATT GTGACTATCA GCATTCTGGT GCAAAATGAAC TTTTCTCCAT CATCGACTGT 420
10 GGAAATTTGA TACTTTTAAA GCATATTTCT CTATGAGCAC AGGTCCTCCT AGTGAACTT 480
AATTTGACAA AGGGTGTCTAT ATGCTTTCCCT AACCTGAWTT GTATTAACAT TCACAGAGCC 540
TACATTTTCT CATTAGGGTT RTGATGCTCA GTATCTTTCC AAGTGCCAGG CAGRGCTTNC 600
15 CTTTCTGAT CAAACATACC ATTTTGTGTA TTTCACTACT ATAGACAGTC ACTTCTGCAG 660
TCCCAATTTA AAAATGCAGA ACTGCTTTAT CCAAGAATGC TGAAAAATAC TGTCTATCC 720
20 AGGTTTCTTA AACTATAAAA GCAGATTTTG CTTTGTGTTG TTAATCATAG GCATGGCCGA 780
GCATGTGGA TTAGCCTGAG GCTTAAATC AGATGCATGT CTGGTAAGAT GACCACTGTC 840
TCATATCAA GAGCCTGCAG AGCCATTTTC CAGACCTGTG ATTGCCCAGA ACACATAGTC 900
25 CCCACGTTTC TAATTTGGAG CAAATCTAAA AG 932

30

(2) INFORMATION FOR SEQ ID NO: 234:

(i) SEQUENCE CHARACTERISTICS:

35

- (A) LENGTH: 2786 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 234:

TTAGCAGGGT GAGCTGTTAA AACAGCACAC ATCTCTCATC CCTCTTCCT TTATTCCCCC 60
CTGGGTTTCA GAAAGGAAG ATATATGGG ACCACCTCCC CCTTCTTTGA TCCAGCATC 120
45 TCAGTCCCCC TCCCAACCCT CCATATGGCT CTCAATGGTG CTCACTTGCT TGAAGCAGG 180
CTCCAATAG GGAGGGGCT GCCCTTACA GTCTCTTTGA CTGTAAGACA GGGCTCTGTA 240
TCAGTGAGAC GATGAGAAAA GTCCCAGGCT AATGGCAGAA ATTTGCACTT TGAACATGTG 300
50 TGTTTTGTG TTGTGGAACC TGAGATTCCT TATTTATTAA CAGGAAGTCT GATTTTTTTT 360
TTTTGGAGTC TTGTGTGCTA TATTTGTGG GGCTGGGAGA GAGAGATTAG ATTATTTTGA 420
55 CATGGGATCC CTCCATAAC AGGTACTTTG AAGGCAAGAC ATAGGGTTGA AGAAGCACAA 480
CCAGCCTCTG AAATCATAGC TCTCCAGTGG CTTTAAAGA AAGCTGGTCC TCAGCACTAA 540
CAAAATCACT ACAATAGCCT AGTGCTTTT TGGAAGCCTT TTTAGGGAAG AATGTTAGGT 600
60

	TCATGGTAAC TAGTATGCTC TTTGAGATTT TTACAGTGTT GAAACTTAAG AATTTTGAGA	660
	GGGTGAGGAG GGTGTTCAG AATCTAAATT ACAGATAGAT GATTGTTTCT TGTGAATTTG	720
5	TTTCTTTTCC TTTTMTTGTG TCCCTACCAT TTCCTTACAT TTCCCTTGGG GCCCATCTCT	780
	GGCTCCTTGC TTTTGTGTTT TGCTTTGCT TTATCAGTTC ATTCCAGCTC CCTGTTAGTG	840
10	AAGGACACTG CTGTTAGTGA AGGAACAAAG TCTATGAGTC CTAATAATTT AAGTCAAAGA	900
	AAACTGCTCT GTTCCCTT TAGTAACACT TCTGAAGAGG AAAAATTCA ATAGCCAAAG	960
	TTAATAATCC TATATAATAA TTGCTTTGGC TTTCACCTAA AATTCTGGGC ATCACAATTT	1020
15	CCTTGGGATA GAGGTGTGTG TGGGAATAG ATTGCTTATT GCTGTTCACT GGAGAGAAAA	1080
	GGTAGTGT TTGTACAAGG TCATACCGCC AGAAGCCCCA AATCCTATTT TGGCTCATCT	1140
20	TCAGGTAAAG AGTAATTCCT ATCCTGTGTG CCTCAGAAGC TAGAATCGAA GGCTTACCCT	1200
	ATTCAITGTT TATTGTGAGA AATGCATGAT GGCTCTTGA AAGAATGACG TTTTGCTGGA	1260
	AAAAAAAAA AGAACAGTTT GTGTTTCACA AACATGGCTT ATCAATTTT TCAAAGAATT	1320
25	CTTTTTCCTC AAAAAGAGGA GTAACAAAT GTCATTTCTG AAAGAGGCTT ACTTTATACC	1380
	AACTAGTGTG AGCATTTGGG ATGCCAGGGA ACAGAGAGTG AGACACCTAC AATCACCAGT	1440
30	CTCAAATGCG CTATTGTTTC TTTTCAGAGT GTTGCAGATT TGCCATTTCT CCATAATATG	1500
	GGGATAGAAA ATGGAATAAA GATAGAAGGG ATGTAGAATA TGCTTTCCTG CCAACATGGT	1560
	TTGGAGTCGA CTTTGGTATA TTGACTAGAT TTGAAAATAC AAGATTGATT AGATGAATCT	1620
35	ACAAAAAAGT TGTCTCTCTC TCAGGTCCCT TTTACACTTT TTGACTAACT AGCATCTATA	1680
	TTCCACACTT AGCTTTTGTG TCACACTTAT CCTTGTCTC CGTAAATTC ATTTGCAGTG	1740
40	GTTAGTCATC AGATATTTTA GCCACCTACA CAAAAGCAAA CTGCATTTT AAAAATCTTT	1800
	CTGAGATGGG AGAAAATGTA TTCTCCTTTC CTATACCGCT CTCCCAACAA AAAAACAAT	1860
	AGTAGTCTCT ACTAATTAGA AACITGCTGT ACTTTTCTT TTCTTTTAGG GGTCAAGGAC	1920
45	CCTCTTTATA GCTACCATTG GCCTACAATA AATTATTGCA GCAGTTTGCA ATACTAAAAT	1980
	ATTTTTTATA GACTTTATAT TTTTCTTTT GATAAAGGGA TGCTGCATAG TAGAGTTGGT	2040
50	GTAATTAAAC TATCTCAGCC GTTCCCTGC TTTCCCTTCT GCTCCATATG CCTCATGTG	2100
	CTTCCAGGGA GCTCTTTTAA TCTTAAAGTT CTACATTTCA TGCTCTTAGT CAAATTCTGT	2160
	TACCTTTTAA ATAACCTCTC CCACTGCATA TTTCCATCTT GAATTGGTGG TTCTAAATTC	2220
55	TGAAACTGTA GTTGAGATAC AGCTATTTAA TATTCTGGG AGATGTGCAT CCTCTTCTT	2280
	KGTGGTGGC CAAGGTGTG TTGCGTAACT GAGACTCCTT GATATGCTTC AGAGAATTTA	2340
60	GGCAACACT GGCCATGGCC GTGGGAGTAC TGGGAGTAAA ATAAAAATAT CGAGGTATAG	2400

ACTAGCATCC ACATAGAGCA CTTGAACCTC CTTTGTACCT GTTTGGGGAA AAAGTATAAT 2460
GAGTGTACTA CCAATCTAAC TAAGATTATT ATAGTCTGGT TGTTTGAAAT ACCATTTTTT 2520
5 TCCTCTTTTG TGTTTTTCCC ACTTTCCAAT GTACTCAAGA AAATTGAACA AATGTAATGG 2580
ATCAAATTAA AATATTTTAT TTCTTAAAAG CCTTTTTTGC CTGTTGTAAT GTGCAGGACC 2640
10 CTCTCCTTT CATGGGAGAG ACAGGTAGTT ACCTGAATAT AGGTTGAAAA GGTATGTAA 2700
AAAGAAATTA TAATAAAGG GATACTTTCG TTTTCAAATC TTTGTTTTCT CTTATCTAG 2760
GTAAGGCATA TTAAAAATAA ATATGT 2786

15

(2) INFORMATION FOR SEQ ID NO: 235:

20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 458 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 235:

GGGTGCAGGA ATTGGGCACG AGAGAATGTT TGATTTTCTT TCCTATTTTA AGGATCTTCT 60
30 CTCTTGTTGA TGTGAAAAC TTACCTTAGT GAAGATGTTT TTCAACATGC TGTGTCCTT 120
TACCTGCATA ATCAGAGCTA TGCATCTATT CAAAGTGATG ATCTGTGGGA TAGTTTAAAT 180
GAGGTCACAA ACCAAACACT AGATGTAAAG AGAATGATGA AAACCTGGAC CCTGCAGAAA 240
35 GGATTTCTCT TAGTGACTGT TCAAAAGAAA GGAAAGGAAC TTTTATACA ACAAGAGAGA 300
TTCTTTTTAA ATATGAAGCC TGAAATTCAG CCTTCAGATA CAAGGTACAT GCCCTCTTTC 360
40 TTTTCATGCC ATCTCTTTTG CACTCTCAGG TGGAAATATT TTTAAGTGT TTATAATCAT 420
AAGTTCTTGT GAAACCTAAC AAGATTATCC CTCCTTAA 458

45

(2) INFORMATION FOR SEQ ID NO: 236:

50 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 591 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 236:

AGGATGAAGA GGAAATTATC TCTTGGATTG CTCTCCAGGA AATCCTTCTC TATACTTTAA 60
AAGCTCTTGT TCTTTCTAG GATCCAATG TGCTGATGTC TGCTAACAGT CAGGATACAA 120

60

TTAAGGTGCT AGAATTGGTA TGAAGGGTTA ACTCAAGTCA AATTGTACTT GATCCTGCTG 180
 AAATACATCT GCAGCTGACA ATGAGAGARG AAACAGAAAA TGTTCATGTA TGTCTCTCCC 240
 5 CAAAGTCATC ATGGGTTTGG GATTGTGTTT GAATATTTT TCTTTTTC TTKTCCCTCC 300
 TTTATGAGCC TTTGGGACAT TGGGAATACC CAGCCAAC TCACCATCA ATGTAAC TCC 360
 ATGGACATTG CTGCTCTTGG TGGTGTATC TAATTTTGT GATAGGAAA CAAATCTTT 420
 10 TGAATAAAAA TAAATAACWA AACAATAAAA GTTTATTGAG CCACAGTTGA GCTTGGAAG 480
 TTTTGTCAA ATGCNGCAAG AGATAACTCT TTTTANGAAG TAGCATATGT GAACTATAAT 540
 15 GTAACAGTGA ATAATTGTA AAGTTCGTAT TTCCCAACCT CTTTGGGAAT T 591

20 (2) INFORMATION FOR SEQ ID NO: 237:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1286 base pairs

(B) TYPE: nucleic acid

25 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 237:

30 TCTTTTAAAG GTACAGCAGG GAAGAACTGG AAACACAGAG AAAGAACTG CCCTTCCATC 60
 TACAAAAGCT GAGTTTACTT CTCCTCCTTC TTTGTTCAAG ACTGGGCTTC CACCGAGCAG 120
 GAGATTACCT GGGCAATTG ATGTTATCGG TCAGACTATA ACTATCAGCC GAGTAGAAGG 180
 35 CAGGCGACGG GCAAATGAGA ACAGCAACAT ACAGGTCCTT TCTGAAAGAT CTGCTACTGA 240
 AGTAGACAAC AATTTTAGCA AACCACCTCC GTTTTTCCTT CCAGGAGCTC CTCCCACTCA 300
 40 CCTTCCACCT CTTCCATTTT TTCCACCTCC TCCGACTGTC AGCACTGCTC CACCTCTGAT 360
 TCCACCACCG GGTTTTCCTC CTCCACCAGG CGCTCCACCT CCATCTCTTA TACCAACAAT 420
 AGAAAGTGGG CATTCCTCTG GTTATGATAG TSGTCTGCA CGTGCACTTC CATATGGCAA 480
 45 TCGGATGAAG AACGATACAG ATACAGGGAA TATGCAGAAA GAGGTTATGA GCGTCACAGA 540
 GCAAGTCGAG AAAANGAAGA ACGACATAGA GAAAGACGAC ACAGGGAGAA AGAGGAAACC 600
 50 AGACATAAGT CTTCTCGAAG TAATAGTAGA CGTCGCCATG AAAGTGAAGA AGGAGATAGT 660
 CACAGGAGAC ACAAACACAA AAAATCTAAA AGAAGCAAAG AAGGAAAAGA AGCGGGCAGT 720
 GAGCCTGCCC CTGAACAGGA GAGCACCAGG GCTACACCTG CAGAATAGGC ATGGTPTTGG 780
 55 CCTTTGTGT ATATTAGTAC CAGAAGTAGA TACTATAAAT CTTGTTATTT TTCTGGATAA 840
 TGTTTAAGAA ATTTACCTTA AATCTTGTTC TGTTTGTTAG TATGAAAAGT TAACPTTTTT 900
 60 TCCAAAATAA AAGAGTGAAT TTTTCATGTT AAGTTAAAAA TCTTTGTCTT GTACTATTTT 960

AAAAATAAAA AGACAGCAAT GACTTTATAT CCAAGAAAGG AATGTGAATG AGTCACTTAA 1020
5 CAGGGAATCT AAAGAGCTGT GTTAGCTGTG TACATACACA GATTATCTGA GAAAAGGTCA 1080
AGGGTTCCAC TTGGGCCACA GTTTTTTTGT TAATCAAACA CCACTCTCTT AAGRGGCTGC 1140
ATCACAARG GCAACCAARG GGCCCTCTT ARGCTTTGA GGATTAAAC TAGTCTTTAT 1200
10 CCATTACTGC TGTGGACACT CTGGCTTRG TATWTTAGG GGGGNTCCTT ACCTTTTTTT 1260
GGTTTCCNC ACCTTTTGG TTGGGC 1286

15

(2) INFORMATION FOR SEQ ID NO: 238:

(i) SEQUENCE CHARACTERISTICS:
20 (A) LENGTH: 734 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 238:

ATGGCAGCGC AGAAGGACCA GCAGAAAGAT GCCGAGCGG AAGGGCTGAG CGGCACGACC 60
CTGCTGCCGA AGCTGATTCC CTCCGGTGCA GGCCGGGAGT GGCTGGAGCG GCGCCGCGCG 120
30 ACCATCCGGC CCTGGAGCAC CTCGTGGAC CAGCAGCGCT TCTCAGGCC CCGCAACCTG 180
GGAGAGCTGT GCCAGCGCCT CGTACGCAAC GTGGAGTACT ACCAGAGCAA CTATGTGTTT 240
35 GTGTCCTGG GCCTCATCCT GTACTGTGTG GTGACGTCCC CTATGTTGCT GTTGGCTCTG 300
GCTGTCTTTT TCGGCGCCTG TTAACATTCT CTATCTGCGC ACCTTGAGT CCAAGCTTGT 360
CCTCTTTGGC CGAAAGGTGA GCCCAGCGCA TCATATGCTC TGGCTGGAGG CATCTCCTTC 420
40 CCCTCTTCTT GGCTGGCTGG TCGGGGCTCG GCCGTCTTCT GGGTGTGGG AGCCACCTG 480
GTGGTCATCG GCTCCACGC TGCCTTCCAC CAGATTGAGG CTGTGGACGG GGAGGAGCTG 540
45 CAGATGGAAC CCGTGTGAGG TGTCTTCTGG GACCTGCCG CCTCCGGGC CAGCTGCCCC 600
ACCCCTGCCC ATGCTGTCC TGCACGGTCT GCTGCTCGG CCCACAGCGC CGTCCCATCA 660
CAAGCCCGGG GAGGGATCCC GCCTTTGAAA ATAAAGCTGT TATGGGTGTC ATTCAAAAAA 720
50 AAAAAAAAAA AAAA 734

55

(2) INFORMATION FOR SEQ ID NO: 239:

(i) SEQUENCE CHARACTERISTICS:
60 (A) LENGTH: 809 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 239:

5 CGGGGTCTTC AGGGTACCGG GCTGGTTACA GCAGCTCTAC CCCTCACGAC GCARACATGG 60
CAGCGCAGAA GGACCAGCAG AAAGATCCCG AGGCGGAAGG GCTGAGCGGC ACGACCCCTGC 120
10 TGCCGAAGCT GATTCCCTCC GGTGCAGGCC GGGAGTGGCT GGAGCGGCGC CGCGCGACCA 180
TCCGCCCTG GAGCACCTTC GTGACCAGC AGCGCTTCTC ACGGCCCCGC AACCTGGGAG 240
AGCTGTGCCA GCGCCTCGTA CGCAACGTGG AGTACTACCA GAGCAACTAT GTGTTCGTGT 300
15 TCCTGGGCCT CATCCTGTAC TGTGTGGTGA CGTCCCCTAT GTTGCTGGTG GCTCTGGCTG 360
TCTTTTTCGG CGCCTGTTAC ATTCTCTATC TGCGCACCTT GGAGTCCAAG CTTGTGCTCT 420
20 TTGGCCGAGA GGTGAGCCCA GCGCATCAGT ATGCTCTGGC TGGAGGCATC TCCTTCCCTT 480
TCTTCTGGCT GGCTGGTGCG GGCTCGGCCG TCTTCTGGGT GCTGGGAGCC ACCCTGGTGG 540
TCATCGGCTC CCACGCTGCC TTCCACCAGA TTGAGGCTGT GGACGGGGAG GAGCTGCAGA 600
25 TGGAACCCGT GTGAGGTGTC TTCTGGGACC TGCCGGCCTC CCGGGCCAGC TGCCCCACCC 660
CTGCCCATGC CTGTCTTGCA CGGCTCTGCT GCTCGGGCCC ACAGCGCCGT CCCATCACAA 720
30 GCCCCGGGAG GGATCCCGCC TTTGAAAATA AAGCTGTTAT GGGTGTCAAT CAGGAAAAAA 780
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA 809

35

(2) INFORMATION FOR SEQ ID NO: 240:

(i) SEQUENCE CHARACTERISTICS:

40 (A) LENGTH: 2201 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 240:

TCGACCCACG CGTCCGGCAA CATGGCGGCT GCCGTGGTGC AGCGCCCGGG CTGAGCGACA 60
GCAAGTGCAG CGGGCTCCTA CCCCAGGTGA GGGGTGGCCT CCGCGTGGGA TCGTGCCCTC 120
50 TTCAGCCCGC TCCTGTCCCC GACATCAGT GTATTCCGCA CGTCCCCTCC GCGCTGTGTG 180
TCTACTGAGA CGGGAGGCG TGACAGGGCC CGGGTCCCTT CTCAGTGGTG CTCTGTGCTT 240
55 CAGGGCAAGC TCCCCGTCTC CGGGCGCACT TCCCTCGCCT GTGTTCGGTC CATCCTCCTT 300
TCTCCAGCCT CTTCCCTCG CAGGCGGATG AMCCGGACGA CGGGCCAGTG CCTGGCACCC 360
CGGGGTGGCC ARGGTCCAMG GGAACCCGA AGTCCGAGGA GCCCARGTC CCGAACCAGG 420
60

	ARGGGCTGCA GCGCATCAMC GGCCTGTCTC CCGGCCGTTC GGCTCTCATA GTGGCGGTGC	480
	TGTGCTACAT CAATCTCCTG AACTACATGG ACCGCTTCAC CGTGGCTGGC GTCCTTCCCG	540
5	ACATCGAGCA GTTCTTCAAC ATCGGGGACA GTAGCTCTGG GCTCATCCAG ACCGTGTTCA	600
	TCTCCAGTTA CATGGTGTG GCACCTGTGT TTGGCTACCT GGGTGACAGG TACAATCGGA	660
10	AGTATCTCAT GTGCGGGGGC ATTGCCTTCT GGTCCCTGGT GACACTGGGG TCATCCTTCA	720
	TCCCCGGAGA GCATTCTGG CTGCTCCTCC TGACCCGGGG CCTGGTGGGG GTCGGGGAGG	780
	CCAGTTATTC CACCATCGCG CCCACTCTCA TTGCCGACCT CTTTGTGGCC GACCAGCGGA	840
15	CCGGATGCTC AGCATCTTCT ACTTTGCCAT TCCGGTGGGC AGTGGTCTGG GCTACATTGC	900
	AGGCTCCAAA GTGAAGGATA TGGCTGGAGA CTGGCACTGG GCTCTGAGGG TGACACCGGG	960
20	TCTAGGAGTG GTGGCCGTC TGCTGCTGTT CCTGGTAGTG CGGGAGCCGC CAAGGGGAGC	1020
	CGTGGAGCGC CACTCAGATT TGCCACCCCT GAACCCACC TCGTGGTGGG CAGATCTGAG	1080
	GGCTCTGGCA AGAAATCCTA GTTTCGTCTT GTCTTCCCTG GGCTTCACTG CTGTGGCCTT	1140
25	TGTCACGGGC TCCCTGGCTC TGTGGGCTCC GGCATTCTTG CTGCGTTCCC GCGTGGTCTT	1200
	TGGGGAGACC CCACCCTGCC TTCCCGGAGA CTCTGTCTCT TCCTCTGACA GTCTCATCTT	1260
30	TGGACTCATC ACCTGCCTGA CCGGAGTCTT GGGTGTGGGC CTGGGTGTGG AGATCAGCCG	1320
	CCGGCTCCGC CACTCCAACC CCCGGGCTGA TCCCTGTGTC TGTGCCACTG GCCTCCTGGG	1380
	CTCTGCACCC TTCTCTTCC TGTCCCTTGC CTGCGCCCGT GGTAGCATCG TGGCCACTTA	1440
35	TATTTTMTAC TTCATTGGAG AGACCTCTCT GTCCATGAAC TGGGCCATCG TGGCCGACAT	1500
	TCTGCTGTAC GTGGTGATCC CTACCCGACG CTCCACCGCC GAGGCCTTCC AGATCGTGCT	1560
40	GTCCACCTG CTGGGTGATG CTGGGAGCCC CTACCTCATT GGCCTGATCT CTGACCGCCT	1620
	GCGCCGGAAC TGGCCCCCTT CCTTCTTGTC CGAGTCCCGG GCTCTGCAGT TCTCGCTCAT	1680
	GCTCTGCGCG TTTGTGTGGG CACTGGGCGG CGCACTTTCC TGGGCACCGC CATCTTCATT	1740
45	GAGGCCGACC GCCGGCGGGC ACAGCTGCAC GTGCAGGGCC TGCTGCACGA AGCAGGGTCC	1800
	ACAGACGACC GGATTGTGGT GCGCCAGCGG GGCCGCTCCA CCCGCGTGCC CGTGGCCAGT	1860
50	GTGCTCATCT GAGARGCTGC CGCTACCTA CCTGCACATC TGCCACAGCT GGCCCTGGGC	1920
	CCACCCACG AAGGGCCTGG GCCTAACCCC TTGGCCTGGC CCAGCTTCCA GAGGGACCCT	1980
	GGGCGGTGTG CCAGCTCCCA GACACTACMT GGGTAGCTCA GGGGAGGAGG TGGGGTCCA	2040
55	GGAGGGGAT CCCTCTCCAC AGGGGCAGCC CCAAGGGCTC GGTGCTATT TTAACGGAAT	2100
	AAAATTGTGA GCCAGACCCC AGGTGCCTGC TCTCGTCTTT CTCTGGGTGG CCTCTGATCT	2160
60	TGCACCCCGT CTTCAACCCA GGGCTCCTGA AGACTGTGGG T	2201

(2) INFORMATION FOR SEQ ID NO: 241:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1661 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

10

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 241:

15	GTCCCTCCCG ACATCGAGCA GTTCTCAAC ATCGGGGACA GTAGCTCTGG GCTCATCCAG	60
	ACCGTGTICA TCTCCAGTTA CATGGTGTG GCACCTGTGT TTGGCTACCT GGGTGACAGG	120
	TACAATCGGA AGTATCTCAT GTGCGGGGGC ATTGCCCTCT GGTCCCTGGT GACACTGGGG	180
20	TCATSCTTCA TCCCCGGAGA GCATTCTTG CTGCTCCTCC TGACCCGGGG CCTGGTGGGG	240
	GTGCGGGAGG CCAGTTATTC CACCATCGCG CCCACTCTCA TTGCCGACCT CTTTGTGGCC	300
25	GACCAGCGGA SCGGATGCTC AGCATCTTCT ACTTTGCCAT TCCGGTGGGC AGTGGTCTGG	360
	GCTACATTGC AGGCTCCAAA GTGAAGGATA TGGCTGGAGA CTGGCACTGG GCTCTGAGGG	420
	TGACACCGGG TCTAGGAGTG GTGGCCGTTT TGCTGCTGTT CCTGGTAGTG CGGGAGCCGC	480
30	CAAGGGGAGC CGTGGAGCGC CACTCAGATT TGCCACCCCT GAACCCACCC TCGTGGTGGG	540
	CAGATYTGAG GGCTCTGGCA AGAAATCCTA GTTTCGTCT GTCTTCCCTG GGCTTCACTG	600
35	CTGTGGCCTT TGTACGGGC TCCCTGGCTC TGTGGGCTCC GGCATTCTTG CTGCGTTCCC	660
	GCGTGGTCTT TGGGGAGACC CCACCTGCC TTCCCGGAGA CTCTGCTCT TCCTCTGACA	720
	GTCTCATCTT TGGACTCATC AACTGCCTGA CCGGAGTCTT GGTGTGGGC CTGGGTGTGG	780
40	AGATCAGCCG CCGGYTCCGC CACTCCAACC CCCGGGCTGA TCCCCTGGTC TGTGCCACTG	840
	GCCTCTGGG CTCTGCACCC TTCTCTTCC TGTCCCTTGC CTGCGCCCGT GGTAGCATCG	900
45	TGGCCACTTA TATTTTCATC TTCAATGGAG AGACCCTCCT GTCCATGAAC TGGGCCATCG	960
	TGGCCGACAT TCTGCTGTAC GTGGTGATCC CTACCCGACG CTCCACCGCC GAGGCCTTCC	1020
	AGATCGTGCT GTCCACCTG CTGGGTGATG CTGGGAGCCC CTACCTCATT GGCCTGATCT	1080
50	CTGACCGCCT GCGCCGGAAC TGCCCCCCT CCTTCTGTG CGAGTTCCGG GCTCTGCAGT	1140
	TCTCGCTCAT GCTCTGCGCG TTTGTGGGG CACTGGGCGG CGCACTTTCC TGGGCACCGN	1200
55	CATCTTCATT GAGGCCGACC GCCGGCGGGC ACAGCTGCAC GTGCAGGGCC TGTGCACGA	1260
	AGCAGGTCC ACAGACGACC GGATTGTGGT GCCCCAGCGG GGCCGCTCCA CCCGCGTGCC	1320
	CGTGGCCAGT GTGCTCATCT GAGAGGCTGC CGCTCACCTA CCTGCACATC TGCCACAGCT	1380
60	KGCCCTGGGC CCACCCACG AAGGGCCTGG GCCTAACCCC TTGGCCTGGC CCAGCTTCCA	1440

GAGGGACCCCT GGGCCGTGTG CCAGCTCCCA GACACTACMT GGGTAGCTCA GGGGAGGAGG 1500
TGGGGGTCCA GGAGGGGAT CCCTCTCCAC AGGGGNCACC CCAAGGGCTC GGTGCTATTT 1560
5 GTAACGGAAT AAAATTGTA GCCAGACCCC AGGTGCCTGC TCTCGTCTTT CTCTGGGTGG 1620
CCTCTGATCT TGCACCCCGT CTTACCCCCA GGGCTCCTGA A 1661

10

(2) INFORMATION FOR SEQ ID NO: 242:

15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1146 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 242:

NGACAGAAAA GCAGAAGATG AGACTCTGTT CATTCACMTT TCCTAGGCCC ATCCTGTGGT 60
25 CATCTTTCCC CCTCCCATCA TACCTCCTCC TTCCTGGAGC CTCTGCCGGC TTGGCTGTAA 120
TGGTGGCACT TACCTGGATA TTTCAGTGGG AGGATGAAAG GCGAGACTCA CCCTACGCGG 180
TGGGACAGAT GGGGAGAGGA AAAAGCAGA GATNGCCAGG AGAGGGGTGC AGGACAAACC 240
30 AGAGAGGTTG GGTCAAGGGA AAAGTGTNGG GAGAAAGTGG GGTGCAGGCC CTGCAGGCCG 300
GTTTAGCCAG CAGCTGCGGC CTCCCAGGGC CCTTGGCATC CAACTTCGCA GACAGGGTAC 360
35 CAGCCTCCTG GTGTGTATCA TAGGATTTGT TCACATAGTG TTATGCATGA TCTTCGTAAG 420
GTTAAGAAGC CGTGGTGGTG CACCATGACA TCCAACCCGT ATATATAAAG ATAAATATAT 480
ATATATATGT ATGTAAATTA TAGCACTGAG GGCCTGCTG CCCTGCTGGA CCAAGCAAAA 540
40 CTAAGCCTTT TGGTTTGGGT ATTATGTTTC GTTTTGITAT TTGTTTGTTT TTGTGGCTTG 600
TCTTATGTCG TGATAGCACA AGTGCCAGTC GGATTGCTCT GTATTACAGA ATAGTGTTTT 660
45 TAATTCATCA ATGTTCTAGT TAATGTCTAC CTCAGCACCT CCTCTTAGCC TAATTTTAGG 720
AGGTTGCCCA ATTTTGTTTC TTCAATTTTA CTGGTTACTT TTTGTACAA ATCAATCTCT 780
TTCTCTCTMT CTCTCCTCCC CACCTCTCAC CCTTGCCCTC TCCATCTCCC TCTCCCGCCC 840
50 TCCCCTCCTC CCTCTGGCTC CCGTCTCAT TTCTGTCCAC TCCATTCTCT CTCCCTCTCT 900
CCTGCCTCCT GCTGCCCCCT CCCAGCCCA CTTSCCCGAG TTGTGCTTGC CGCTCCTTAT 960
55 CTGTTCTAGT TCCGAAGCAG TTCACTCGA AGTTGTGCAG TCCTGGTTGC AGCTTTCCGC 1020
ATCTGCCTTC GTTTCGTGTA GATTGACGCG TTTCTTTGTA ATTTCACTGT TTCTGACAAG 1080
ATTTAAAAAA AAAAAAAGGA AAAAAAATA AAAAAAATC TCGAGGGGGG GCCCGGTACC 1140
60

CAATTG

1146

5

(2) INFORMATION FOR SEQ ID NO: 243:

(i) SEQUENCE CHARACTERISTICS:

10

(A) LENGTH: 1350 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 243:

20	AACCCACGGC TGCTGCGGCA GGGCGTGGAG GGCAGAGGGC CGCGGAGGCG CAGTTGCAAA	60
	CATGGCTCAG AGCAGAGACG GCGGAAACCC GTTCGCCGAG CCCAGCGAGC TTGACAACCC	120
25	CTTTCAGCCA CCACCAGCCT ATGAGCCTCC AGCCCCCTGCC CCATTGCCTC CACCCTCAGC	180
	TCCCTCCTTG CAGCCCTCGA GAAAGCTCAG CCCACAGAA CCTAAGAACT ATGGCTCATA	240
	CAGCACTCAG GCCTCAGCTG CAGCAGCCAC AGCTGAGCTG CTGAAGAAAC AGGAGGAGCT	300
	CAACCGGAAG GCAGAGGAGT TGGACCGAAG GAGNCGAGAG CTGCAGCATG CTGCCCTGGG	360
	RGGCAGAGCT ACTCGACAGA ACAATGGGCC CCTCTACCT TCTTTTGTG CAGTTCAGCC	420
30	CTGCTTTTTC CAGGACATCT CCATGGAGAT CCCCCAAGAA TTTCAGAAGA CTGTATCCAC	480
	CATGTACTAC CTCTGGATGT GCAGCAGSST GGCTCTTCTC CTGAACCTCC TCGCCTGCCT	540
	GGCCAGCTTC TGTGTGGAAG CCAACAATGG CGCAGGCTTT GGGCTTTCTA TCCTCTGGGT	600
35	CCTCCTTTTC ACTCCCTGCT CCTTTGTCTG CTGGTACCGC CCCATGTATA AGGCTTTCCG	660
	GAGTGACAGT TCATTCAATT TCTTCGTTT CTCTTCATT TTCTTCGTCC AGGATGTGCT	720
40	CTTTGTCTC CAGGCCATTG GTATCCCAGG TTGGGGATTG AGTGGCTGGA TCTCTGCTCT	780
	GGTGGTGCCG AAGGCAACAC AGCAGTATCC GTGCTCATGC TGCTGGTCGC CCTGCTCTTC	840
	ACTGGCATTG CTGTGCTAGG AATTGTCATG CTGAAACGGA TCCACTCCTT ATACCGCCGC	900
45	ACAGGTGCCA GCTTTCAGAA GGCCAGCAA GAATTGCTG CTGGTGTCTT CTCCAACCCT	960
	GCGGTGCGAA CCGCARCTTG CCAATGCAGC CGCTGGGGCT GCTGAAAATG CCTCCGGGC	1020
50	CCCCTGACCC CTGACTGGGA TGCCCTGGCC CTGCTACTTG AGGGAGCTGA CTTAGCTCCC	1080
	GTCCCTAAGG TCTCTGGGAC TTGGAGAGAC ATCACTAACT GATGGCTCCT CCGTAGTGCT	1140
55	CCCAATCCTA TGGCCATGAC TGCTGAACCT GACAGGCGTG TGGGGAGTTC ACTGTGACCT	1200
	AGTCCCCCA TCAGGCCACA CTGCTGCCAC CTCTCACACG CCCCCACCA GCTTCCCTCT	1260
	GCTGTGCCAC GGCTGTGTCT TCGGTATTAT AAATAAAAG AAAGTGAAC TGGAAAAAA	1320
60	AAAAAAAAA AAAAAAAAAAG GGGGNCNC	1350

5 (2) INFORMATION FOR SEQ ID NO: 244:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 1529 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 244:

15 TCCCAGAGGC CGGGGGGTC CAGCTCTGCC TG TAGCAGAG CCCTGAGGAG GAGGAGGAAG 60
AGGATGTGCT GAAATACGTC CGGAGATCT TTTTCAGCTA GGCATAAAC TGTGCACTGA 120
ACTGTCTGCC GAGAGCAGCT GGAGGACAGC TGAGCTTCCA CTGGTGCTGC TGGGCCGMCC 180
20 GCCTGTGGGA ATGGGGCTCT CTGTGCTCCT ACCTTTGTGC CTTCITGGGC CTGGCAGATT 240
CACCTCAGGC CAGAAGCCCC TGGACACTCC GGGCCTTGGG GTGCCGTTCT GAGTGTGCGG 300
25 AAGGCAGGAC TCAAAATGAG ATCCCATTTG ACTCCCTCTG TATGTACTGT GCCCTCTCCT 360
GGCTCTTGAG GCTCTGGAGT CCCAATTGTC TGTGTTAGTC AGTGACCAGG TTCCAGGGAA 420
AATRATGTCA TGTGGTGGTC CAACTTACTG GAACCAAAGA GACAGTACTT TGCAAAGAAA 480
30 AGGATCACTG CCAGGTGCAC TGAATTGCT ACAGTTTAGT CCGCATGATC TCTCCTGAAG 540
GAGGAAGCCT GTTTCAAAAA TAGTTTCCAT CATGAGTCTA TCAATGAGCT CCCACCTCTC 600
35 CAGCCAGCCT AGAAAGCAAA CGAGCTGCCC ACAGTTCTCT GCCCTGTCTG GGAGGTTGAG 660
GCCACAGTGT ATAGACTGGT AAGCCAGACA GGCTCTCTCC CGCAAGCTGC TACCTTGCTT 720
TCACCTGTAC CTTGGTCCCC GGGCAGCTAG CTATAAAGCA AGAGGGACAG GAGCCCAGAA 780
40 GAGACACTGA GGACAAGAGA TCACACCAGA GTACATGTCT CTGCCTCTGT TTTCAAGTGT 840
GCTTTGGACA GGAATATATG AATAAATCAC TGCCATACAG GTTTTCCAAT ACACAAGTGC 900
45 TAGAAAATAC ACACAATTCC CCAATGCGTA AGTTGTGCTA ATGTCTTTCC AAGTTCTGGG 960
TTGGGAAGTG GAGGGTGGCA GCGTTTGTTT GTGCGCAACC GTCCAGTCCT GTTCACAGCG 1020
AGGATTTGGA GTCTCCAGG GTCTCATCAT GGGAGTGATT TGTCAGCGGA CGCCTCTGCC 1080
50 CTGTCTGGCT TCAGGTCCAG GGAAGCTTTG AAGCAGTCAA GCCTGTCTTT TGTACCCCAT 1140
GTGTCTGTG TTTGTTGAGT CACTCAGAGA TCACTCCTGG ACCTCTGGGG TTGGAGTTCC 1200
55 AGTGATGGCT TATGGCGGCC CACTCACTAT GGTGGGCTGA GTGGAAGCTC CTTAACCATG 1260
TCCCCAGAGA CACTGAGGTG CTCGCTCTTT TAATGTCTTC GTTTGTTGCC GTAAGTTCTT 1320
TGCTAGGTTT CATTTTGCA TTTGGCAAAT CAGCCTGGAA GTCTGGCCCC ATGACAGCAA 1380
60

TCACTCCCTC CCCACCTCC TGAAGCTAGA GGAAGATTG CTCAGATCCA TTAATTAAAG 1440
CAGGAATTGG TGTGACAATG AGCTGCATGG TTTAGGGAGT CTTGGGAGC CTGGAAGTC 1500
5 CTGAAGGACA AACAATCTTG TACTAAGAA 1529

10 (2) INFORMATION FOR SEQ ID NO: 245:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1537 base pairs
(B) TYPE: nucleic acid
15 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 245:

20 GTGCGAGGTC CCCGCCAGCC CCCAGCGGCC TTCCCGGCCC GGGGCGCTCC CAGAGCAAAC 60
GAGGCCCTG AGAGCTCCAC CTAGTTCACA GGATAAAATC CCACAGCAGA ACTCGGAGTC 120
AGCAATGGCT AAGCCCCAGG TGGTTGTAGC TCCTGTATTA ATGTCTAAGC TGTCTGTGAA 180
25 TGCCCCTGAA TTTTACCCTT CAGGTTATTC TTCCAGTTAC ACAGAATCCT ATGAGGATGG 240
TTGTGAGGAT TATCCTACTC TATCAGAATA TGTTCAAGAT TTTTGAATC ATCTTACAGA 300
30 GCAGCCTGGC AGTTTGAAG CTGAAATGA ACAGTTTGCA GAGACCCTGA ATGGTTGTGT 360
TACAACAGAT GATGCTTTGC AAGAACTTGT GGAATCATC TATCAACAGG CCACATCTAT 420
CCCCAAATTC TCTTATATGG GAGCTCGCCT GTGTAATTAC CTGTCCCATC ATCTGACAAT 480
35 TAGCCCCAG AGTGCCAAT TCCGCCAATT GCTACTTCAA AGATGTCCGA CTGAATATGA 540
AGTTAAAGAT CAAGCTGCAA AAGGGGATGA AGTTACTCGA AAACGATTTC ATGCATTTGT 600
40 ACTCTTTCTG GGAGAACTTT ATCTTAACCT GGAGATCAAG GGAACAAATG GACAGGTTAC 660
AAGAGCAGAT ATTCTTCAGG TTGGTCTTCG AGAATTGCTG AATGCCCTGT TTTCTAATCC 720
TATGGATGAC AATTTAATTT GTGCAGTAAA ATTGTTAAAG TTGACAGGAT CAGTTTGGGA 780
45 AGATGCTTGG AAGGAAAAAG GAAAGATGGA TATGAAGAA ATTATTGAGA GAATTGAAAA 840
CGTTGTCTTA GATGCAAACT GCAGTAGAGA TGTAACACAG ATGCTCTTGA AGCTTGTAGA 900
50 ACTCCGGTCA AGTAACGGG GCAGAGTCCA TGCAACTTCA ACATATAGAG AAGCAACACC 960
AGAAAATGAT CCTAACTACT TTATGAATGA ACCAACATTT TATACATCTG ATGGTGTTC 1020
TTTCACTGCA GCTGATCCAG ATTACCAAGA GAAATACCAA GAATTACTTG AAAGAGAGGA 1080
55 CTTTTTTCCT GATTATGAAG AAAATGGAAC AGATTTATCC GGGGCTGGTG ATCCATACTT 1140
GGATGATATT GATGATGAGA TGGACCCAGA GATAGAAGAA GCTTATGAAA AGTTTGTGTT 1200
60 GGAATCAGAG CGTAAGCGAA AACAGTAAAG TTAAATTTCA GCATATCAGT TTTATAAAGC 1260

AGTTTAGGTA TGGTGATTTA GCAGAACACA AGAGAGCAAG AAAATGTGTC ACATCTATAC 1320
CAAATTRAGG ATGTTGAGTT ATGTTACTAA TGTATGCAAC TTAAATTTTG TTAAACACTA 1380
5 TCTGCCAAAA TAAACTTTAT TOCCTATAAC TTA AAAATGTG TATATATATA TAATAGTTTA 1440
TTATGTACAG TTAATCTAC TGTTTGGCT GCAATAAAAT CGATTTTGAA ATAAAWRAAA 1500
10 AAAAAAAAAA AAGGGNGGCC GCTCTAGAGG ANCCAAG 1537

15 (2) INFORMATION FOR SEQ ID NO: 246:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 506 base pairs
(B) TYPE: nucleic acid
20 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 246:

25 TGCAGGATTT GGCAGGACC CSCCGCGTG GCGGTGCTA TCGCTTCGCA GAACCTACTC 60
AGGCAGCCAG CTGAGAAGAG TTGAGGAAA GTGCTGCTGC TGGGTCTGCA GACCGCATGG 120
ATAACGTGCA GCCGAAAATA AAACATCGCC CCTTCTGCTT CAGTGTGAAA GGCCACGTGA 180
30 AGATGCTGCG GCTGGATATT ATCAACTCAC TGGTAACAAC AGTATTCATG CTCATCGTAT 240
CTGTGTTGGC ACTGATACCA GAAACCACAA CATTGACAGT TGGTGGAGGG GTGTTTGAC 300
35 TTGTGACAGC AGTATGCTGT CTGCGGACG GGGCCCTTAT TTACCGGAAG CTTCTGTTC 360
ATCCAGCGG TCCTTACCAG AAAAGCCTG TGCATGAAA AAAAGAAGTT TTGTAATTTT 420
ATATTACTTT TTAGTTTGAT ACTAAGTATT AAACATATTT CTGKATTATT CCAAAAAAAAA 480
40 AAAAAAAAAA AAAAAAATT TGGTGG 506

45

(2) INFORMATION FOR SEQ ID NO: 247:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1348 base pairs
(B) TYPE: nucleic acid
50 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 247:

55 GTCTTCTTT TNCTGTTTG AGTTGGTGAG TGAGTGAATA GGGTAACATG GGCCTTCAGG 60
ATGACCCCTT GGAAGTGTG CGAGTTCCTT AAATCTCAGC TGGGATCCTG GACCTGGGAG 120
60 GCCCTGTGA GGGCCAGCTC TGGAAAAACC TGGGAGTTGA TGCCGAGGY TGGGAAGAAC 180

	TCCTGCTCGAG GGCAGGGTGC CCTGGAACAC TGGTAGTTCT GGGGCTGGGA GGGAGAGGGG	240
5	CTCCGGCTTT CTCTGAAATG AACACTGCTC TTCAGCAGTT CAAGTACTTG TTCTCAAAAC	300
	ATTTTCTAAT TGATTGGTAG GTTTCATAA GCATTGTTTC TTTAAGGCAT GGAAAGGGAA	360
	GAATGCTCAA GCAAGTCATG TTTGTTTTCA GTGGGATGGG CCCGCGTTCT CACTGCTGGG	420
10	GGCTTCCCCCT TGCATGTGGC ACCTTTGTGC AGGGCCACCA GGCAGACTCT TCCCACCTTC	480
	TCCCAGTGAA GCACCAAGGG GCTTGAACCG TAATTTGGCT AATCAGAGGC ATTTTPTTTTG	540
15	TCCTAGTATC TTTCACACTT GTCCAACCGT CTTATTTTTC TAAAAGTTCT GTTGCTTGTA	600
	TTAACACGAA ACTAGAGAGA AATAGTTTCT GAAGCCAGTT TATTGTGAAG ATCCCCAAGG	660
	GGAGGTTCCG TAGAGAAAA TAGTAAGCTG GTTTAGAAAC TGACGAGGGC AAACAGCCAG	720
20	GACGCATTGG AGAGGAATTT GCCAAAGATC TACCCTGAGA TAACGCCTGT CCAGTGTCTT	780
	CACCACGTGA ATAACCAGCG CTCCAAAGTG TTTTCTGCT TTGAAAAAAA AAATTCACA	840
25	AGCTTTTAAA GGTGCATTTA AGAATCCATG TGACTTTAGA ATGGAAGTGC CGGCCCTGGC	900
	AACTGTCACG TGTGCTAGAA GGTTCGATGC CTCTGGAATG CATGTGATAC TCATCTCCAT	960
	TTTGTTCCTT TGATTGCATT TTTGTTCTTT TAGCAGATCT GTCCCTGTGG GTGGTGTCTA	1020
30	AGAAGTCGGA CACCTTGGTT TTTGTGTTAG ATTGAGCTGG GCAGCTGCAA TCAGCTTCTT	1080
	TATATGCAA TTAGGCACGA CCCATCTGTG GTTCCCTGGT TGGTGGCTAA TGAAGTGAGG	1140
35	GGAGGGAGGG ATGTCACCCC AAAAGTAGGC CCTCCCATTG GCTTTGGCCA GGCCAGACAC	1200
	TTACATCGT TTACATGGTT CTGTGTAATT TTAAAGTTTA TGTGTATAAA GCGAAGCTGT	1260
	TTCTGTGAAA CTGTATATTT TGTAAATAAA TATATTGCTA CTTTGAGAWR AAAAAAAAAA	1320
40	AAAAACTCGA GGGGGGCCCG GTACCCAA	1348

45 (2) INFORMATION FOR SEQ ID NO: 248:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 1766 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 248:

55	GTGCCGAATC GGCAGAGCG CACGAGCGC CACGAGACA GCGGAGTAA AGGGACTTGA	60
	GCGAGCCAGT TGCCGGATTA TTCTATTTCC CCTCCCTCTC TCCGCCCCG TATCTCTTTT	120
60	CACCTTCTC CCACCTCGC TCGGTASCA TGGCGGAGCG TCGGCGGCCA CTCAGTCCCA	180

	TTCCATCTCC TCGTCGTCTT TCGGAGCCGA GCCGTCCGCG CCCGGCGGCG GCGGGAGCCC	240
	AGGAGCCTGC CCCGCCCTGG GGACGAAGAG CTGCAGCTCC TCCTGTGCGG TGCACGATCT	300
5	GATTTTCTGG AGAGATGTGA AGAAGACTGG GTTTGTCTTT GGCACCACGC TGATCATGCT	360
	GCTTTCCTTG GCAGCTTTCA GTGTCATCAG TGTGGTTTCT TACCTCATCC TGGCTCTTCT	420
10	CTCTGTCACC ATCAGCTTCA GGATCTACAA GTCCGTCATC CAAGCTGTAC AGAAGTCAGA	480
	AGAAGGCCAT CCATTCAAAG CCTACCTGGA CGTAGACATT ACTCTGTCTT CAGAAGCTTT	540
	CCATAATTAC ATGAATGCTG CCATGGTGCA CATCAACAGG GCCCTGAAAC TCATTATTCTG	600
15	TCTCTTTCTG GTAGAAGATC TGGTTGACTC CTTGAAGCTG GCTGTCTTCA TGTGGCTGAT	660
	GACCTATGTT GGTGCTGTTT TTAACGGAAT CACCCTTCTA ATTCTGTCTG AACTGCTCAT	720
20	TTTCAGTGTC CCGATTGTCT ATGAGAAGTA CAAGACCCAG ATTGATCACT ATGTTGGCAT	780
	CGCCCCGAGAT CAGACCAAGT CAATTGTTGA AAAGATCCAA GCAAAACTCC CTGGAATCGC	840
	CAAAAAAAG GCAGAATAAG TACATGGAAA CCAGAAATGC AACAGTTACT AAAACACCAT	900
25	TTAATAGTTA TAACGTCGTT ACTTGTTACTA TGAAGGAAAA TACTCAGTGT CAGCTTGAGC	960
	CTGCATTCCA AGCTTTTPTT TTAATTGTTT GTTTTCTCCC ATCCTTTCCC TTTAACCTTC	1020
30	AGTATCAAGC ACAAAAATTG ATGGACTGAT AAAAGAACTA TCTTAGAACT CAGAAGAAGA	1080
	AAGAATCAAA TTCATAGGAT AAGTCAATAC CTTAATGGTG GTAGAGCCTT TACCTGTAGC	1140
	TTGAAAGGGG AAAGATTGGA GGTAAAGAGG AAAATGAAAG AACACCTCTG GGTCTTCTG	1200
35	TCCAGTTTTC AGCACTAGTC TTACTCAGCT ATCCATTATA GTTTTGCCCT TAAGAAGTCA	1260
	TGATTAACTT ATGAAAAAAT TATTTGGGGA CAGGAGTGTG ATACCTTCCT TGGTTTTTTT	1320
40	TTGCAGCCCT CAAATCCTAT CTTCCTGCCC CACAATGTGA GCAGCTACCC CTGATACTCC	1380
	TTTTCTTTAA TGATTTAACT ATCAACTTGA TAAATAACTT ATAGGTGATA GTGATAATTC	1440
	CTGATTCCAA GAATGCCATC TGATAAAAA GAATAGAAAT GGAAAGTGGG ACTGAGAGGG	1500
45	AGTCAGCAGG CATGCTGCGG TGGCGGTCAC TCCCTCTGCC ACTATCCCA GGAAGGAAA	1560
	RGCTCCGCCA TTTGGGAAAG TGGTTTCTAC GTCAGTGGAC ACCGGTTCTG AGCATTAGTT	1620
50	TGAGAACTCG TTCCCGAATG TGCTTTCTCT CCTCTCCCT GCCACCTCA AGTTTAATAA	1680
	ATAAGGTGT ACTTTTCTTA CTATAAAATA AAAAAAAAAA AACTCGAGGG GGGCCCGGTA	1740
	CCCAAATCGC CGGATATGAT CGTAAA	1766

55

(2) INFORMATION FOR SEQ ID NO: 249:

60

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2664 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 249:

	AGTGTCTCTCG GAGCAGGCGG AGTAAAGGGA CTTGAGCGAG CCAGTTGCCG GATTATTCTA	60
10	TTTCCCCTCC CTCTCTCCCG CCCCGTATCT CTTTTCACCC TTCTCCCACC CTCGCTCGCG	120
	TASCATGGCG GAGCGTCGGC GGCCACTCAG TCCCATTCOA TCTCCTCGTC GTCCTTCGGA	180
15	GCCGAGCCGT CCGCGCCCGG CGGCGGCGGG AGCCCAGGAG CCTGCCCCGC CCTGGGGACG	240
	AAGAGCTGCA GCTCCTCCTG TGGGTGCAC GATCTGATTT TCTGGAGAGA TGTGAAGAAG	300
	ACTGGGTTTG TCTTGGCAC CACGCTGATC ATGCTGCTTT CCCTGGCAGC TTTCAGTGTC	360
20	ATCAGTGTGG TTTCTTACCT CATCCTGGCT CTTCTCTCTG TCACCATCAG CTTCAGGATC	420
	TACAAGTCCG TCATCCAAGC TGTACAGAAG TCAGAAGAAG GCCATCCATT CAAAGCCTAC	480
25	CTGGACGTAG ACATTACTCT GTCTCAGAA GCTTTCATA ATTACATGAA TGCTGCCATG	540
	GTGCACATCA ACAGGGCCCT GAAACTCATT ATTCTCTCTT TCTGGTAGA AGATCTGGTT	600
	GACTCCTTGA AGCTGGCTGT CTTCAATGTT CTGATGACCT ATGTTGGTGC TGTTTTTAAC	660
30	GGAATCACCC TTCTAATCTT TGCTGAACGT CTCATTTTCA GTGTCCCGAT TGTCTATGAG	720
	AAGTACAAGA CCCAGATTGA TCACTATGTT GGCATCGCCC GAGATCAGAC CAAGTCAATT	780
35	GTGAAAAGA TCCAAGCAAA ACTCCCTGGA ATCGCCAAAA AAAAGGCAGA ATAAGTACAT	840
	GGAAACCAGA AATGCAACAG TTAATAAAC ACCATTTAAT AGTTATAACG TCGTTACTTG	900
	TACTATGAAG GAAAATACTC AGTGTGAGCT TGAGCCTGCA TTCCAAGCTT TTTTTTTAAT	960
40	TTGGTGTTTT CTCCCATCCT TTCCCTTTAA CCTCAGTAT CAAGCACAAA AATTGATGGA	1020
	CTGATAAAAG AACTATCTTA GAACTCAGAA GAAGAAAGAA TCAAAATTCAT AGGATAAGTC	1080
45	AATACCTTAA TGGTGGTAGA GCCTTTACCT GTAGCTTGAA AGGGGAAAGA TTGGAGGTAA	1140
	GAGAGAAAAT GAAAGAACAC CTCTGGGTCC TTCTGTCCAG TTTTCAGCAC TAGTCTTACT	1200
	CAGCTATCCA TTATAGTTTT GCCCTTAAGA AGTCATGATT AACTTATGAA AAAATTATTT	1260
50	GGGGACAGGA GTGTGATACC TTCCTTGGTT TTTTTTTCGA GCCCTCAAAT CCTATCTTCC	1320
	TGCCCCACAA TGTGAGCAGC TACCCCTGAT ACTCCTTTTC TTTAATGATT TAACTATCAA	1380
55	CTTGATAAAT AACTTATAGG TGATAGTGAT AATTCTGAT TCCAAGAATG CCATCTGATA	1440
	AAAAAGAATA GAAATGGAAA GTGGGACTGA GAGGGAGTCA GCAGGCATGC TGCGGTGGCG	1500
	GTCACCTCCCT CTGCCACTAT CCCAGGGAA GGAAARGCTC CGCCATTTCG GAAAGTGGTT	1560
60	TCTACGTAC TGGACACCGG TTCTGAGCAT TAGTTTGAGA ACTCGTTCCC GAATGTGCTT	1620

TCCTCCCTCT CCCCTGCCCA CCTCAAGTTT AATAAATAAG GTTGTACTTT TCTTACTATA 1680
 AAATAAATGT CTGTAAGTGC TGTGCACTGC TGTAAACTTG TTAGAGAAAA AAATAACCTG 1740
 5 CATGTGGGCT CCTCAGTTAT TGAGTTTPTG TGATCCTATC TCAGTCTGGG GGGGAACATT 1800
 CTCAAGAGGT GAAATACAGA AAGCCTTTTT TTCTTGATCT TTTCCCGAGA TTCAAATCTC 1860
 10 CGATTCCCAT TTGGGGGCAA GTTTTTTTCT TCACCTTCAA TATGAGAATT CAGCGAAGTT 1920
 GAAAGAAAAA TCATCTGTGA GTTCCTTCAG GTTCTCACTC ATAGTCATGA TCCTTCAGAG 1980
 GGAATATGCA CTGGCGAGTT TAAAGTAAGG GCTATGATAT TTGATGGTCC CAAAGTACGG 2040
 15 CAGCTGCAAA AAGTAGTGGA AGGAAATTGT CTACGTGTCT TGGAAAAATT AGTTAGGAAT 2100
 TTGGATGGGT AAAAGGTACC CTTGCCTTAC TCCATCTTAT TTTCTTAGCC CCCTTTGAGT 2160
 20 GTTTTAACTG GTTTCATGTC CTAGTAGGAA GTGCATTCTC CATCCTCATC CTCTGCCCTC 2220
 CCAGGAAGTC AGTGATTGTC TTTTGGGCT TCCCCTCCAA AGGACCTTCT GCAGTGAAG 2280
 TGCCACATCC AGTCTTTTC TTTTGTGCT GCTGTGTTA GATAATTGAA GAGATCTTTG 2340
 25 TGCCACACAG GATTTTTTTT TTTTAAAGA AAAACCTATA GATGAAAAAT TACTAATGAA 2400
 ACTGTGTGTA CGTGTCTGTG CGTGCAACAT AAAAATACAG TAGCACCTAA GGAGCTTGAA 2460
 30 TCTTGGTTCC TGTAAATTT CAAATTGATG TGGTATTAAT AAAAAAAAAA AAAACAMAAA 2520
 AAAAAAAAAA AAAAGGGCGG CCGCTCTAGA GGATCCAAGC TTACGTACGC GTGCATGCGA 2580
 CGTCCATAGC TCTTTCTATA GGGGTCCCCC AAATTCCATT CANCGGGCCG TCGGTTTTAN 2640
 35 AAAGGTCGTG ANTGGGGGAA ANCC 2664

40

(2) INFORMATION FOR SEQ ID NO: 250:

(i) SEQUENCE CHARACTERISTICS:

45

- (A) LENGTH: 865 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 250:

CGTGGGAGTG AGGTACCAGA TTCAGCCCAT TTGCCCCGA CGCCTCTKTT CTCGGAATCC 60
 GGGTGTGCG GATTGAGGTC CCGGTTCTTA ACGGTGGGAT CGGTGTCCTC GGGATGAGAT 120
 55 TTGGCGTTTC CTCGGGCTTT TGGTGGGATC GGTGTCCTCA GGATGAGATT TAGGGTTTCC 180
 TCGGGGCTTT CGGGATCTTC ACCTAATATC CGGACTGCAA GATGGAGGAA GCGGGAACC 240
 TAGGAGGCCT GATTAARATG GTCCATCTAC TGGTCTGTG AGGTGCCTGG GGCATGCAAA 300
 60

TGTGGGTGAC CTTCTGTCTCA GGCTTCCTGC TTTTCCGAAG CCTTCCCCGA CATACCTTCG 360
GACTAGTGCA GAGCAAATC TTCCCTTCT ACTTCCACAT CTCCATGGGC TGTGCCITCA 420
5 TCAACCTCTG CATCTTGGCT TCACAGCATG CTTGGGCTCA GCTCACATTC TGGGAGGCCA 480
GCCAGCTTTA CTTGCTGTTC CTGAGCCTTA CGCTGGCCAC TGTCAACGCC CGCTGGCTGG 540
10 AACCCCGCAC CACAGCTGCC ATGTGGGCCC TGCAAACCGT GGAGAAGGAG CGAGGCCTGG 600
GTGGGGAGGT ACCAGGCAGC CACCAGGTC CCGATCCCTA CCGCCAGCTG CGAGAGAAGG 660
ACCCCAAGTA CAGTGTCTTC CGCCAGAATT TCTTCCGCTA CCATGGGCTG TCCTCTCTTT 720
15 GCAATCTGGG CTGCGTCTG AGCAATGGGC TCTGTCTCGC TGGCCTTGCC CTGGAAATAA 780
GGAGCCTCTA GCATGGGCCC TGCATGCTAA TAAATGCTTC TTCAGAAAAA AAAAAAAAAA 840
AAACTCGAGG GGGGCCCGGT ACCCA 865
20

(2) INFORMATION FOR SEQ ID NO: 251:

25

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2082 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 251:

TGGGGGGGN AATGGGTGTC TGGCTCANGG ATTGCCNAAT CTGGAAATTC TCCATAACTT 60
35 GCTAGCTTGT TTTTMTTMTT TTTTMTTACA CCCCCCGCC CCACCCCGG ACTTGACAAA 120
TGTTCATGA TCTCAGCAGA GTTCTTCATG TGAAACGTTG ATCACCTTTG AAGCCTGCAT 180
40 CATTCACATA TTTTMTCTTC TTCTTCCCT TCAGTTCATG AACTGGTGT CATTTTCTGT 240
GTGTGTGTGT GTTTTATTTT GTTTGGATTT TTTTMTTAA TTTTACTTTT AGAGCTTGCT 300
GTGTGCCCCA CCTTTTMTCC AACCTCCACC CTCCTCCTT CTCACCCAT CTCTCCGAG 360
45 ATGAAAGAAA AAAAAAGCA AAGTTTMTT TTCTTCTCCT GAGTTCTTCA TGTGAGATG 420
AGCTTGCAAA GAAAAAATA ATGTGAAATG TTATAGACTT GCAGCGTGCC GAGTCCATC 480
50 GGGTTTMTT TTTAGCATG TTATGCTAAA ATAGAGAAA AAATGCTCAT GAACCTTCCA 540
CAATCAAGCC TGCATCAACC TTCTGGGTGT GACTTGTGAG TTTTGGCCTT GTGATGCCAA 600
ATCTGAGAGT TTAGTCTGCC ATTAATAAAA CTCATTCTCA TCTCATGCAT TATTATGCTT 660
55 GCTACTTGT CTTAGCAACA ATGAACTATA ACTGTTTCAA AGACTTTATG GAAAAGAGAC 720
ATTATATTAA TAAAAAATA AAGCCTGCAT GCTGGACATG TATGGTATAA TTATTTTTC 780
60 CTTTTTTTTT CCTTTTGGCT TGGAAATGGA CGTTCGAAGA CTTATAGCAT GGCATTCATA 840

CTTTTGTTTT ATTGCCTCAT GACTTTTTTG AGTTTAGAAC AAAACAGTGC AACCGTAGAG 900
CCTTCTTCCC ATGAAATTTT GCATCTGCTC CAAAACTGCT TTGAGTTACT CAGAACTTCA 960
5 ACCTCCCAAT GCACTGAAGG CATTCCTTGT GCAAAGATAC CAGAATGGGT TACACATTTA 1020
ACCTGGCAAA CATTGAAGAA CTCITRATGT TTTCTTTTAA ATAAGAATGA CGCCCCACTT 1080
10 TGGGGACTAA AATTGTGCTA TTGCCGAGAA GCAGTCTAAA ATTTATTTTT TAAAAAGAGA 1140
AACTGCCCCA TTATTTTGGG TTTGTTTTAT TTTATTTTAA TATTTTGGG CTTTGGTCA 1200
TTGTCAAATG TGAATGCTC TGGGTCTCTA GTATATAATT TAATCTAGT TTTTATAATC 1260
15 TGTAGCCCA GTTAAATGT ATGCTACAGA TAAAGGAATG TTATAGATAA ATTTGAAAGA 1320
GTTAGGTCTG TTTAGCTGTA GATTTTPTAA ACGATTGATG CACTAAATTG TTTACTATTG 1380
20 TGATGTTAAG GGGGTAGAG TTTGCAAGGG GACTGTTTAA AAAAGTAGC TTATACAGCA 1440
TGTGCTTGCA ACTTAAATAT AAGTTGGGTA TGTGTAGTCT TTGCTATACC ACTGACTGTA 1500
TTGAAAACCA AAGTATTAAG AGGGGAAACG CCCCTGTTA TATCTGTAGG GGTATTTTAC 1560
25 ATTCAAAAAT GTATGTTTTT TTTCTTTTC AAAATTAAAG TATTTGGGAC TGAATTGCAC 1620
TAAGATATAA CCTGCAAGCA TATAATACAA AAAAAATTG CAAACTGTT TAGAACGCTA 1680
30 ATAAATTTA TGCAGTTATA AAAATGGCAT TACTGCACAG TTTTAAGATG ATGCAGATTT 1740
TTTACAGTT GTATTGTGGT GCAGAACTGG ATTTCTGTA ACTTAAAAAA AAATCCACAG 1800
TTTAAAGGC AATAATCAGT AAATGTTATT TTCAGGGACT GACATCCTGT CTTTAAAAAG 1860
35 AAATGAAAAG TAAATCTTAC CACAATAAAT ATAAAAAAT CTGTTCAGTT ACTTTTCTTT 1920
TACATATTTT GCTGTGCAAA ATTGTTTAT ATCTTGAGTT ACTAACTAAC CACGCGTGT 1980
40 GTTCCTATGT GCTTTTCTTT CATTTTCAAT TCTGGTTATA TCAAGAAAAG AATAATCTAC 2040
AATAATAAAC GGCATTTTTT TTTGAAAAAA AAAAAAAAAA AA 2082

45

(2) INFORMATION FOR SEQ ID NO: 252:

(i) SEQUENCE CHARACTERISTICS:
50 (A) LENGTH: 1482 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 252:

CAGGCAGGCT GGCCCCGGGG ACTTCTCTCT GGCCCTGCTC CCTCCGAGCG CTCGCGGTT 60
60 GCGCGCCTGG CCCCTACGGA GTCCTTAGCC AGGATGGAGG CTGTTGTGAA CTTGTACCAA 120

GAGGTGATGA AGCACGCAGA TCCCCGGATC CAGGGCTACC CTCTGATGGG GTCCCCCTTG 180
CTAATGACCT CCATTCTCCT GACCTACGTG TACTTCGTTC TCTCACTTGG GCCTCGCATC 240
5 ATGGCTAATC GGAAGCCCTT CCAGCTCCGT GGCTTCATGA TTGTCTACAA CTTCTCACTG 300
GTGGCACTCT CCCTCTACAT TGTCTATGAG TTCCTGATGT CGGGCTGGCT GAGCACCTAT 360
ACCTGGCGCT GTGACCCTGT GGAATATTCC AACAGCCCTG AGGCACCTAG GATGGTTCCG 420
10 GTGGCCTGGC TCTTCCTCTT CTCCAAGTTC APTGAGCTGA TGGACACAGT GATCTTTATT 480
CTCCGAAAGA AAGACGGGCA GGTGACCTTC CTACATGTCT TCCATCACTC TGTGCTTCCC 540
15 TGGAGCTGGT GGTGGGGGGT AAAGATTGCC CCGGGAGGAA TGGGCTCTTT CCATGCCATG 600
ATAAACTCTT CCGTGCATGT CATAATGTAC CTGTACTACG GATTATCTGC CTTTGGCCCT 660
GTGGCACAAC CCTACCTTTG GTGGAAAAAG CACATGACAG CCATTGAGCT GATCCAGTTT 720
20 GTCTGGTCT CACTGCACAT CTCCCAGTAC TACTTTATGT CCAGCTGTAA CTACCAGTAC 780
CCAGTCATTA TTCACCTCAT CTGGATGTAT GGCACCATCT TCTCATGCT GTTCTCCAAC 840
25 TTCTGGTATC ACTCTTATAC CAAGGGCAAG CGGCTGCCCC GTGCACCTCA GCAAAATGGA 900
GCTCCAGGTA TTGCCAAGGT CAAGGCCAAC TGAGAAGCAT GGCCTAGATA GGCGCCACC 960
TAAGTGCTC AGGACTGCAC CTTAGGGCAG TGTCCGTCAG TGCCCTCTCC ACCTACACCT 1020
30 GTGACCAAGG CTTATGTGGT CAGGACTGAG CAGGGGACTG GCCCTCCCCT CCCACAGCT 1080
GCTCTACAGG GACCACGGCT TTGGTTCTTC ACCCACTTCC CCGGGGCAGC TCCAGGGATG 1140
35 TGGCCTCATT GCTGCTGCC ACTCCAGAGC TGGGGGCTAA AAGGGCTGTA CAGTTATTTC 1200
CCCCTCCCTG CCTTAAACT TGGGAGAGGA GCACTCAGGG CTGGCCCCAC AAAGGGTCTC 1260
GTGGCCTTTT TCCTCACACA GAAGAGGTCA GCAATAATGT CACTGTGGAC CCAGTCTCAC 1320
40 TCCTCCACCC CACACACTGA AGCAGTAGCT TCTGGGCCAA AGGTCAGGGT GGGCGGGGCG 1380
CTGGGAATAC AGCCTGTGGA GGCTGCTTAC TCAACTTGTG TCTTAATTAA AAGTGACAGA 1440
45 GGAAACCAAA AAAAAAAAAA AAAAATCGA GGGGGGCCG TA 1482

50 (2) INFORMATION FOR SEQ ID NO: 253:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 834 base pairs

(B) TYPE: nucleic acid

55 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 253:

60 GGCACGAGCG CCGTTGCCCG CCTGGCCCT ACGAGTCTT TAGCCAGGAT GGAGGCTGTT 60

GTGAACTTGT ACCAAGAGGT GATGAAGCAC GCAGATCCCC GGATCCAGGG CTACCCCTCTG 120
ATGGGGTCCC CCTTGCTAAT GACCTCCATT CTCCTGACCT ACGTGTACTT CGTTCTCTCA 180
5 CTTGGGCCTC GCATCATGGC TAATCGGAAG CCCTTCCAGC TCCGTGGCTT CATGATTGTC 240
TACAACTTCT CACTGGTGGC ACTCTCCCTC TACATGTCTT ATGAGTTCCT GATGTCGGGC 300
10 TGGCTGAGCA CCTATACCTG GCGCTGTGAC CCTCAGGACT GCACCTTAGG GCAGTGTCCG 360
TCAGTGCCTT CTCCAMCTAC ACCTGTGACC AAGGCTTATG TGGTCAGGAC TGAGCAGGGG 420
ACTGGCCCTC CCCTCCCCAC AGCTGCTCTA CAGGGACCAC GGCTTTGGTT CCTCACCAC 480
15 TTCCCCCGGG CAGCTCCAGG GATGTGGCCT CATTGCTGTC TGCCACTCCA GAGCTGGGGG 540
CTAAAAGGGC TGTACAGTTA TTTCCCCCTC CCTGCCTTAA AACTTGGGAG AGGAGCACTC 600
20 AGGGCTGGCC CCACAAAGGG TCTCGTGGCC TTTTCTCTCA CACAGAAGAG GTCAGCAATA 660
ATGTCACTGT GGACCCAGTC TCACTCTCTC ACCCCACACA CTGAAGCAGT AGCTTCTGGG 720
CCAAAGGTCA GGGTGGGCGG GGGCCTGGGA ATACAGCCTG TGGAGGCTGC TTA CTCAACT 780
25 TGTGCTCTTA TTAAGGTGA CAGAGGAAAC CACGAAAAA AAAAAAAAAA AAAA 834

30

(2) INFORMATION FOR SEQ ID NO: 254:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1508 base pairs
35 (B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 254:

TTGAACTTTT AAAATTTTAG ATCAGCAAAC TCTAAGATCC TAGAATGGAA GCTGTTCTCTC 60
ATTTCTCCAT GCTCACCTC CCAGGTCAGC GAGATGGTGA AGAAGCTGCA CGCGCAACA 120
45 CCACCAACGT TCGGAGTGA CCTCATCAAT GAGCTTGTGG AGAACTTTGG CAGATGTCCC 180
AAGTGGTCTG GTCGGCAAGC CTTGTCTTTT GTCTGCCAGA CTGTCAATGA GGATGACTGC 240
CTTCCCATGG ACCAGTTTGC TGTGCATCTC ATGCCGCATC TGCTAACCTT AGCAAATGAC 300
50 AGGGTTCCCTA ACGTCCGAGT GCTGCTTGCA AAGACATTAA GACAACTCT ACTAGAAAAA 360
GACTATTCTT TGGCCTCTGC CAGCTGCCAC CAGGAGGCTG TGGAGCAGAC CATCATGGCT 420
55 CTTCAGATGG ACCGTGACAG CGATGTCAAG TATTTTGCAA GCATCCACCC TGCCAGTACC 480
AAAATCTCCG AAGATGCCAT GAGCACAGCG TCCTCAACCT ACTAGAAGGC TTGAATCTCG 540
GTGTCTTTCC TGCTTCCATG AGAGCCGAGG TTCAGTGGGC ATTGCCACG CATGTGACCT 600
60

GGGATAGCTT TCGGGGAGG AGAGACCTTC CTCTCCTGCG GACTTCATTG CAGGTGCAAG 660
TTGCCTACAC CCAATACCAG GGATTTCAG AGTCAAGAGA AAGTACAGTA AACACTATTA 720
5 TCTTATCTTG ACTTTAAGG GAAATAATTT CTCAGAGGAT TATAATTGTC ACCGAAGCCT 780
TAAATCCTTC TGTCTTCCTG ACTGAATGAA ACTTGAATTG GCAGAGCATT TTCCTTATGG 840
10 AAGGGATGAG ATTCCCAGAG ACCTGCATTG CTTTCTCCTG GTTTTATTTA ACAATCGACA 900
AATGAAATTC TTACAGCCTG AAGGCAGACG TGTGCCCAGA TGTGAAAGAG ACCTTCAGTA 960
TCAGCCCTAA CTCTTCTCTC CCAGGAAGGA CTTGCTGGGC TCTGTGGCCA GCTGTCCAGC 1020
15 CCAGCCCTGT GTGTGAATCG TTTGTGACGT GTGCAAATGG GAAAGGAGGG GTTTTACAT 1080
CTCCTAAAGG ACCTGATGCC AACACAAGTA GGATTGACTT AAACCTTTAA GCGCAGCATA 1140
TTGCTGTACA CATTTACAGA ATGGTTGCTG AGTGTCTGTG TCTGATTTTT TCATGCTGGT 1200
20 CATGACCTGA AGGAAATTTA TTAGACGTAT AATGTATGTC TGGTGTTTTT AACTTGATCA 1260
TGATCAGCTC TGAGGTGCAA CTTCTTCACA TACTGTACAT ACCTGTGACC ACTCTTGGGA 1320
25 GTGTGCACT CTTTAATCAT GCTGTTTAAA CTGTTGTGGC ACAAGTTCTC TTGTCCAAAT 1380
AAAATTTATT AATAAGATCT ATAGAGAGAG ATATATACAC TTTTGATTGT TTTCTAGATG 1440
TCTACCAATA AATGCAATTT GTGACCTGTA TTAACAAAAA NTAAAAAAC TCGAGGGGGG 1500
30 CCCGGTAC 1508

35

(2) INFORMATION FOR SEQ ID NO: 255:

(i) SEQUENCE CHARACTERISTICS:

40

- (A) LENGTH: 2514 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 255:

45

GAGAGACTCA CACTTCTTTT CCATTATCAC TGACGATGTA GTGGACATAG CAGGGGAAGA 60
GCACCTACCT GTGTGGTGA GGTGTTGTTGA TGAATCTCAT AACCTAAGAG AGGAATTTAT 120
50 AGGCTTCCTG CCTTATGAAG CCGATGCAGA AATTTTGGCT GTGAAATTTT CACTATGAT 180
AACTGAGAAG TGGGATTAA ATATGGAGTA TTGTCGTGGC CAGGCTTACA TTGWTCTAG 240
TGGATTTTCT TCCAAAATGA AAGTTGTTGC TTCTAGACTT TTAGAGAAAT ATCCCCAAGC 300
55 TATCTACACA CTCTGCTCTT CCTGTGCCTT AAATATGTGG TTGGCAAAAT CAGTACCTGT 360
TATGGGAGTA TCTGTTGCAT TAGGAACAAT TGAGGAAGTT TGTCTTTTTT TCCATCGATC 420
60 ACCACAACCT CTTTATAGAAC TTGACAACGT AATTTCTGTT CTTTTTCAGA ACAGTAAAGA 480

	AAGGGGTAAA GAACTGAAGG AAATCTGCCA TTCTCAGTGG ACAGGCAGGC ATGATGCTTT	540
5	TGAAATTTTA GTGGAAGTCC TGCAAGCACT TGTTTTATGT TTAGATGGTA TAAATAGTGA	600
	CACAAATATT AGATGGAATA ACTATATAGC TGGCCGAGCA TTTGTACTCT GCAGTGCAGT	660
	GTCAGATTTT GATTTCAITG TTAATATTGT TGTCTTAA AATGTCCTAT CTTTTACAAG	720
10	AGCCTTTGGG AAAAACCTCC AGGGGCAAAC CTCTGATGTC TTCTTTGCGG CCGGTAGCTT	780
	GACTGCAGTA CTGCATTAC TCAACGAAGT GATTGGAAAA TATTGAAGTT TATCATGAAT	840
15	TTTGGTTTGA GGAAGCCACA AATTGGCAA CCAAAGTTGA TATTCAAATG AAATCCCTG	900
	GGAAATTCGG CAGAGCTCAC CAGGGTAACT TGAATCTCA GCTAACCTCT GAGAGTTACT	960
	ATAAGAAAC CCTAAGTGC CCAACAGTGG AGCACATTAT TCAGGAAGTT AAAGATATAT	1020
20	TCTCAGAACA GCACCTCAA GCTCTTAAAT GCTTATCTCT GGTACCTCA GTCATGGAC	1080
	AATCAAAAT CAATACGTCG GAGGAACACC ATGCTGACAT GTATAGAAGT GACTTACCCA	1140
	ATCTGACAC GCTGTGAGT GAGCTTCATT GTTGAGAAAT CAAATGGAAA CACAGGGGGA	1200
25	AAGATATAGA GCTTCCGTCC ACCATCTATG AAGCCCTCCA CCTGCCTGAC ATCAAGTTTT	1260
	TTCTTAATGT GTATGCATG CTGAAGGTCC TGTGTATTCT TCCTGTGATG AAGGTTGAGA	1320
30	ATGAGCGGTA TGAAATGGA CGAAGCGTC TTAAAGCATA TTTGAGGAAC ACTTTGACAG	1380
	ACCAAAGTCC AAGTAACITG GCTTTGCTTA ACATAAATTT TGATATAAAA CACGACCTGG	1440
35	ATTTAATGGT GGACACATAT ATTAACTCT ATACAAGTAA GTCAGAGCTT CCTACAGATA	1500
	ATTCCGAAAC TGTGGAAAAT ACCTAAGAGA CTTTAAAAA TAGGCTTTCT TATATTGAT	1560
	ATTTGGAAGA AAAAGCCGTA AGTGTATGTA GACCACTTAA TCACTAAATA TCTTTGCCTA	1620
40	TAGGACTCCA TTGAATACAT TAGCCATTGA TAATCTACCT GTTTAAATGG CCCCTGTTTG	1680
	AATCTCAAG CTTTGAAGAC CTACCTGTTT TTCCAGAAGA GAACGTTGAA AGTGCCATGT	1740
	TTCTTTTGC GTGATCTCTG TTGATGGCAC TCTGGAATTG TTTCAGTTAA GTCATTTTAG	1800
45	ACATAGCATT TATTATCACT GTGGATCTCT ACTTGTGGG TGTATGAAT TCTTTGAAGA	1860
	AATATATTTT GAAGAGGTGT GGGAGGAAG AATACATTTT ATAAAATGTT GTAGTGAAGC	1920
50	CCACAATGA CCTTTGACTA ATAGGAGTTT TAAGTATGTT AAAAATCTAT ACTGGACAGT	1980
	TACAAGAAAT TACCGGAGAA AAGCTTGTGA GCTCACCAA CAAGGATTTT AGTGTAGATT	2040
	TTGTCTTTCT TGAAGTTAAA GAAACAAATG ACAAAGTTT AATGGAAAAG CCTGCTGTTG	2100
55	TTCCACATCT CGTTGCTGTT TACATTCCTT TGTGGAGCCT ACATCTTCCT AAGCTTTTTA	2160
	GCAGGTATAT GTTGAACACT TCTGTTTCAT GGTGAGACA GAATCAGAGG CCATGGATAC	2220
60	TGACAACGA TTTGTCTGTT TTTTCTCTCT GTCTTTTCC ATGACTCTTA TATACTGCCT	2280

5 CATCTTGATT TATAAGCAAA ACCTGAAAA CCTACAAAAT AAGTGTTGTG GTTATCTAG 2340
AAAAATATGG AAAATATGTC TGTTATTTTT GGTGAAGAAA ATCAATTTTG TATAGTTTAT 2400
TTCAATCTAA ATAAAATGTG AATTTGTGTT AAAGCTTAGG CACATPATTT TTTGTGGGGT 2460
CAAAACATTC TTGTGTAAAT TCTCTTAAAC ATTTGATAAA CAGCTTCACA ATTC 2514

10

(2) INFORMATION FOR SEQ ID NO: 256:

15

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2357 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 256:

CTGCCTTATG AAGCCGATGC AGAAATTTTG GCTGTGAAAT TTCACACTAT GATAACTGAG 60
25 AAGTGGGGAT TAAATATGGA GTATTGTCGT GGCCAGGCTT ACATTGTCCTC TAGTGGATTT 120
TCTTCCAAAA TGAAAGTTGT TGCTTCTAGA CTTTITAGAGA AATATCCCCA AGCTATCTAC 180
ACACTCTGCT CTTCCTGTGC CTTAAATATG TGGTTGGCAA AATCAGTACC TGTATGGGA 240
30 GTATCTGTTG CATTAGGAAC AATTGAGGAA GTTTGTCTCT TTTTCCATCG ATCACCACAA 300
CTGCTTTTAG AACTTGACAA CGTAATPYCT GTTCTTTTTC AGAACAGTAA AGAAAGGGGT 360
35 AAAGAACTGA AGGAAATCTG CCATTCTCAG TGGACAGGCA GGCATGATGC TTTTGAAATT 420
TTAGTGAAC TCCTGCAAGC ACTTGTTTTA TGTITAGATG GTATAAATAG TGACACAAAT 480
ATTAGATGGA ATAACTATAT AGCTGGCCGA GCATTTGTAC TCTGCAGTGC AGTGTGAGAT 540
40 TTTGATTTCA TTGTTACTAT TGTITGTTCT AAAAATGTCC TATCTTTTAC AAGAGCCTTT 600
GGGAAAAACC TCCAGGGGCA AACCTCTGAT GTCTTCTTTG CGGCCGGTAG CTTGACTGCA 660
45 GTACTGCATT CACTCAACGA AGTGANTGGA AAATATTGAA GTTTATCATG AATTTTGGTT 720
TGAGGAAGCC ACAAATTGG CAACCAAACT TGATATTCAA ATGAAACTCC CTGGGAAATT 780
CCGCAGAGCT CACCAGGGTA ACTTGAATC TCAGCTAACC TCTGAGAGTT ACTATAAAGA 840
50 AACCTAAGT GTCCCAACAG TGGAGCACAT TATTCAGGAA CTTAAAGATA TATTCTCAGA 900
ACAGCACCTC AAAGCTCTTA AATGCTTATC TCTGGTACCC TCAGTCATGG GACAACTCAA 960
55 ATCAATACG TCGGAGGAAC ACCATGCTGA CATGTATAGA AGTGACTTAC CCAATCCTGA 1020
CACGCTGTCA GCTGAGCTTC ATTGTTGGAG AATCAAATGG AAACACAGGG GGAAAGATAT 1080
AGAGCTTCCG TCCACCATCT ATGAAGCCCT CCACCTGCCT GACATCAAGT TTTTTCCTAA 1140
60

TGTGTATGCA TTGCTGAAGG TCCTGTGTAT TCTTCCTGTG ATGAAGGTTG AGAATGAGCG 1200
 GTATGAAAAT GGACGAAAGC GTCTTAAAGC ATATTTGAGG AACACTTTGA CAGACCAAAG 1260
 5 GTCAAGTAAC TTGGCTTTGC TTAACATAAA TTTTGATATA AAACACGACC TGGATTTAAT 1320
 GGTGGACACA TATATTAAAC TCTATACAAG TAAGTCAGAG CTTCTACAG ATAATTCCGA 1380
 10 AACTGTGGAA AATACCTAAG AGACTTTTAA AAATAGCCTT TCTTATATTT GATATTTGGA 1440
 AGAAAAAGCC GTAAGTGTAT GTAGACCACT TAATCACTAA ATATCTTTGC CTATAGGACT 1500
 CCATTGAATA CATTAGCCAT TGATAATCTA CCTGTTTAAA TGGCCCCCTGT TTGAACTCTC 1560
 15 AAGCTTTGAA GACCTACCTG TTCTTCCAGA AGAGAACGTT GAAAGTGCCA TGTTCCTTT 1620
 TGCCTGATCT CTGTTGATGG CACTCTGGAA TTGTTTCAGT TAAGTCATTT TAGACATAGC 1680
 ATTTATTATC ACTGTGGATC TCTACTTGTT GGGTGTATG AATTCTTTGA AGAAATATAT 1740
 20 TTTGAAGAGG TGTGGGAGGA AGGAATACAT TTTATAAAAT GTGTAGTGA AGCCCAAT 1800
 TGACCTTTGA CTAATAGGAG TTTTAAGTAT GTTAAAAATC TATACTGGAC AGTTACAAGA 1860
 25 AATTACCGGA GAAAAGCTTG TGAGCTCACC AAACAAGGAT TTCAGTGTAG ATTTTGTCTT 1920
 TCTTGAACCT AAAGAAACAA ATGACAAAGT TTGAATGGAA AAGCCTGCTG TTGTTCCACA 1980
 TCTCGTTGCT GTTTACATTC CTTTGTGGAG CCTACATCTT CCTAAGCTTT TTAGCAGGTA 2040
 30 TATGTTGAAC ACTTCTGTTT CATGGTTGAG ACAGAATCAG AGGCCATGGA TACTGACAAC 2100
 TGATTTGCT GTTTTTTTC TCTGCTTTT TCCATGACTC TTATATACTG CCTCATCTTG 2160
 35 ATTTATAAGC AAAACCTGGA AAACCTACAA AATAAGTGT GTGGTTTATC TAGAAAAATA 2220
 TGGAAATAT TGCTGTTATT TTTGGTGAAG AAAATCAATT TTGTATAGTT TATTTCAATC 2280
 40 TAAATAAAAT GTGAATTTG TTTAAAGCTT AGGCACATTA TTTTTGTGG GGTCAAAACA 2340
 TTCTGTGTA AATTCTC 2357

45

(2) INFORMATION FOR SEQ ID NO: 257:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 689 base pairs
 50 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 257:

ACTTCTGGT GCAAAAAGAT GTTCAAGCCT TATTTTATAC TTGCCTGCC CTTTCTCTTT 60
 CATTTATTGG AGTGAGCTGC AGCTCTAAGA AGACCTGTTT TTTTGAATGG AGAGTAGCAT 120
 60 CAGGAACCAG GATGTGGGTG CGAGGCGTGC TCCTGGCTGT TGCAGATTGC TGCACCCGGG 180

AGCTCTTAGT GGACAGAGCT AGAGGATATG TGCACGTACT TCCATCTCTC TCTCTGTCTC 240
 5 CGATTTTAGC CCAGCACCAC AGGGTACGTT CCAGTTTTTC TCTCTTTCCA TAGCTGTAAG 300
 GCCCTTTCTG GGAATGGTTC TCATTCTCCT TAATCTATTA TTGGGTCAGT TTTCCCTGCAT 360
 GTCCCCAGCC TCCCATCACT GCCACCCACT CCCACAGAG ATGCCCTGCT CATCCGACTG 420
 10 GGGCTTTGAC TCCCACACTG TGTACCCCTC TTGTGTGGAC GCCCTGCTGC CAAAACCTTC 480
 AGCAACAGC TTTCCAAATG GAAGTTGTCA CTGTCARGGS CTTTACAATC AGCAACAGCA 540
 15 AAATCTACAT GCTGCTGAGG GTCCTGCCTC ATTAAGATGC AATAAATATG TAAGTACATA 600
 AAAACAGCAA TAGAAGAAAC GTAATGCTTT ATTCTCAAAT ATGNATGTCT ACATAGAAAA 660
 GCCAAAATTA TTAAGAATAG TAAGGAATT 689
 20

(2) INFORMATION FOR SEQ ID NO: 258:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2377 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 258:

TCGACCCACG CGTCCGCCGA TGTGATGATT CCTGCGTATT CCAAGAACCG GGCCTATGCC 60
 35 ATCTCTTCA TAGCTCTCAC TGTGATAGGG GACGCCCCCG GCGCTGTGCT ATCCTGTGCC 120
 GGCCACCCIT GCGTTGGTTT TGCTGCTGTA CTGGTGGCGC CCCTGACCGT GGCTGTCTCC 180
 40 TCTTGAAGGA AGCCTGTTC TGATGAACCT GCTGACAGCC ATCATCTACA GTCAGTTCCG 240
 GGGCTACCTG ATGAAATCTC TCCAGACCTC GCTGTTTCGG AGGCGGGTGG GAACCCGGCT 300
 GCCTTGAAG TCCTATCCTC CATGGTGGG GAGGGAGGAG CCTTCCCTCA GGCAGTTGGG 360
 45 GTGAAGCCCC AGAACTTGCT GCAGGTGCTT CAGAAGGTCC AGCTGGACAG CTCCCACAGA 420
 CAGGCCATGA TGGAGAAGGT GCGTTCCTAT GGCACTGTTT TGCTCTCAGC TGAGGAGTTT 480
 50 CAGAAGCTCT TCAACGAGCT TGACAGAAGT GTGGTTAAAG AGCACCCGCC GAGGCCCGAG 540
 TACCAGTCTC CGTTTCTGCA GAGCGNCCCA GTTCTCTTC GGCCACTNAC TACTTTGACT 600
 ACCTGGGGAA CCTCATCGCC CTGGCAAACC TGGTGCCAT TTGCGTGTTC CTGGTGCTGG 660
 55 ATGCAGATGT TGCTGCCTGC TGAGCGTGAT GACTTCATCC TGGGGGTCT CAACTGCGTC 720
 TTCAITGTGT ACTACCTGTT GGAGATGCTG GCTCAAGGTC TTTTGCCCTG GGGCCTGCGA 780
 60 RGGTACYKKT CCTAACCCCA RCAAMGTGTT TTGAACGGG TCCTCAMCGT TTGTCTGGC 840

TGGWWKKGSM GATCTCAACT CTGGCTGTGT ACCGATTGCC ACACCCAGGC TGGAGGCCGG 900
ANATGGTGGG CCTGCTGTCTG CTGTGGGACA TGACCCGCAT ACTGAACATG CTCATCGTGT 960
5 TCCGCTTCCT GCGTATCATC CCCAGCATGA AGCCGATGGC CGTGGTGGCC AGTACCGTCC 1020
TGGGCCTGGT GCAAAACATG CGTGCGTTTG GCGGGATCCT GGTGGTGGTC TACTACGTAT 1080
10 TTGCCATCAT TGGGATCAAC TTGTTTAGAG GCGTCATGTG GGCTCTTCCT GGAAACAGCA 1140
GCCTGGCCCC TGCCAATAGG TCGGCGCCCT GTGGGAGCTT CGAGCAGCTG GAGTACTGGG 1200
CCAACAACCT CGATGACTTT GCGGCTGCCC TGGTCACTCT GTGGAACCTG ATGGTGGTGA 1260
15 ACAACTGGCA GGTGTTTCTG GATGCATATC GGCCTACTA AGGCCCGTGG TCCAAGATCT 1320
ATTTTGTATT GTGGTGGCTG GTGTCGTCTG TCATCTGGGT CAACCTGTTT CTGGCCCTGA 1380
TTCTGGAGAA CTTCCTTCAC AAGTGGGACC CCCGCAGCCA CCTGCAGCCC CTGTCTGGGA 1440
20 CCCCAGAGGC CACCTACCAG ATGACTGTGG AGCTCCTGTT CAGGGATATT CTGGAGGAGC 1500
CCGGGGAGGA TGAGCTCACA GAGAGGCTGA GCCAGCACCC GCACCTGTGG CTGTGCAGGT 1560
25 GACGTCCGGG TCTGCCATCC CAGCAGGGG GGCAGGAGAG AGAGGCTGGC ATAACACAGG 1620
TGCCCATCAT GGAAGAGGCG GCCATGCTGT GGCCAGCCAG GCAGGAAGAG ACCTTTCCTC 1680
TGACGGACCA CTAAGCTGGG GACAGGAACC AAGTCCTTTG CGTGTGGCCC AACAAACATT 1740
30 TACAGAACAG CTGCTGTGTC TTCAGGGAGG CGCCGTGCCC TCCGCTTTCT TTTATAGCTG 1800
CTTCAGTGAG AATTCCCTTG TCGACTCCAC AGGGACCTTT CAGACAAAAA TGCAAGAAGC 1860
35 AGCGGCCTCC CCTGTCCCTT GCAGCTTCCG TGGTGCCTTT GCTGCCGGCA GCCCTTGGGG 1920
ACCACAGGCC TGACCAGGCC CTGCACAGGT TAACCGTCAG ACTTCCGGGG CATTCAGCTG 1980
GGAATGATAC TAATACCTCC GATTTTAGCC CAGCACCACA GGTACGTTT CAGTTTTTAT 2040
40 TTCTTTCCAT AGCTGTAAGG CCCTTTCTGG GAAATGGTTAT CATTCCTCTT AATCTATTAT 2100
TGGGTCAGTT TTCTGCATG TCCCAGCCT CCCATCACTG CCACCCACTC CCCACAGAGA 2160
45 TGCCCTGCTC ATCCGACTGG GGCTTTGACT CCCACACTGT GTACCCCTCT TGTGTGGACG 2220
CCCTGCTGCC AAAACCTTCA GCAAACAGCT TTCCAATGG AAGTTGTAC TGTGAGGGCC 2280
TTTACAATCA GCAACAGCAA AATCTACATG CTGCTGAGGG TCCTGCCTCA TTAAGATGCA 2340
50 ATAAATATGT AAGTACATAA AAAAAAAAAA AAAAAA 2377

55

(2) INFORMATION FOR SEQ ID NO: 259:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1193 base pairs

(B) TYPE: nucleic acid

60

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 259:

5 TCTGNTCGCC GTCGCCCCGC CCTGGCCCTT TGCCCGGTCG GGCGGGACTT CCTGTGTCTGT 60
ATTTCCAAGG ACTCCAAAGC GAGGCCGGGG ACTGAAGGTG TGGGTGTCTGA GCCCTCTGGC 120
10 AGAGGGTTAA CCTGGGTCAA ATGCACGGAT TCTCACCTCG TACAGTTACG CTCTCCCGCG 180
GCACGTCCGC GAGGMYTTGA AGTCCTGAGC GCTCAAGTTT GTCCGTAGTC GAGAGAAGGC 240
CATGGAGGTG CCGCCACCGG CACCGCGGAG CTTCCTCTGT AGAGCATTGT GCCTATTTCC 300
15 CCGAGTCTTT GCTGCCGAAG CTGTGACTGC CGATTCGGAA GTCTTGAGG AGCGTCAGAA 360
GCGGCTTCCC TACGTCCAG AGCCCTATTA CCCGGAATCT GGATGGGACC GCCTCCGGGA 420
20 GCTGTTTGGC AAAGACACAG TGAACACTAG TCTGAATGTA TACCGAAATA AAGATGCCTT 480
AAGCCATTTT GTAATTGCAG GAGCTGTCAC GGAAGTCTT TTAGGATAA ACGTAGGCCT 540
GCGTGGCTGG TGGCTGGTGG CATAATTGGA GCCTTGCTGG GCACTCCTGT AGGAGGCCTG 600
25 CTGATGGCAT TTCAGAAGTA CTCTGGTGAG ACTGTTTCTG AAAGAAAACA GAAGGATCGA 660
AAGGCACTCC ATGAGCTAAA ACTGGAAGAG TGGAAAGGCA GACTACAAGT TACTGAGCAC 720
30 CTCCCTGAGA AAATTGAAAG TAGTTTACAG GAAGATGAAC CTGAGAATGA TGCTAAGAAA 780
ATTGAAGCAC TGCTAAACCT TCCTAGAAAC CCTTCAGTAA TAGATAAACA AGACAAGGAC 840
TGAAAGTGCT CTGAACTTGA AACTCACTGG AGAGCTGAAG GGAGCTGCCA TGTCGGATGA 900
35 ATGCCAACAG ACAGGCCACT CTTTGGTCAG CCTGCTGACA AATTTAAGTG CTGGTACCTG 960
TGGTGGCAGT GGCTTGCTCT TGTCTTTTTC TTTTCTTTTT AACTAAGAAT GGGGCTGTTG 1020
40 TACTTCTACT TTACTTATCC TTAAATTAA ATACATACTT ATGTTTGTAT TAATCTATCA 1080
ATATATGCAT ACATGAATAT ATCCACCCAC CTAGATTTTA AGCAGTAAAT AAAACATTTT 1140
GCAAAAGATT AAAGTTGAAT TTTACAGTTA AAAAAAAAAA AAAAAAAAAA AAA 1193
45

(2) INFORMATION FOR SEQ ID NO: 260:

50

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1262 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
55 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 260:

60 GAAAAACCCA AAGATGCAGA CAATCTCTTT GAACATGAAT TGGGGGCTCT CAATATGGCT

	GCATTACTAC GAAAAGAAGA AAGAGCAAGT CTTCTTAGTA ATCTTGGCCC ATGTTGTAAG	120
	GCGTTGTGCT TCAGACGGGA TTCTGCAATT CGAAAGCAGC TTGTTAAAAA TGAGAAGGGC	180
5	ACCATAAAAC AAGCTTACAC GAGTSCCTCA ATGGTAGACA ATGAATTACT TCGATTGAGT	240
	CTTCGGTTAT TTAAGCGGAA GACTACTTGC CATGCTCCAG GACATGAAAA GACTGAAGAT	300
10	AATAAACTTT CACAGTCCAG TATCCAACAG GAACTGTGTG TGTCTTAAGA CCGAAGTTCA	360
	ATATGGTATT TTTGGTACTG TCTTCCTTCA GCAGTGCATA TTCTTTTGCA AAGTTCTTTG	420
	GTTTGACAAG CATTAGTGAC AAAGGCAGAA AAGATTTATC AGCCATGCTA AAAGAGTGAA	480
15	GAATTTTGAT CTTTAGAGAC ACTAGTTTGT GCCAACTTAA GATTTTACGT TAATTTTAC	540
	ATAGTATTTG ACACTCATGC AAAATAATGT GAAAACATCT AGATTTAGTA GTTTATTCTG	600
	CGCCTTTTGT TAAAACTGAA GATTTTGAA AATGGTTGTC ACTGCTCTTC CAGCCTATGA	660
20	ATATTTTGT GAAATGGAAC CATGGATTTA TGTCTGGATC ATCCATACAG AACCAACAAT	720
	TTTATTCAAA AACAAATGTGT TCATCAAAGT AATTGCTCAC ATTGTGCAGT ACTATGTTGT	780
25	ACAGACCACG TGAAAGGGAA TGCTGGTCTA GCTGGCGTGG TATGTTTATA GGCGAATTTT	840
	AGCAGAAGGA AGCCAAAATA GTTTTTTCCT TTTGAAAGTT TTTTAAAAAT TATTTTCATGG	900
30	GTCTTTTTTT TAATTAATAT GTGTGCATTG TTACAATGTA TGTGGATGT CTTTGTACCC	960
	TAAATGCTTT TTTTGTATC AGAGATTGTG TACTATTTTT ATTTTAAATA AATGTATCTT	1020
	CCCTTTCCCT GTTTTAGATT TACTTTGCTC TTCGTTAATC TTATTCCTGA TGATCTAGAA	1080
35	CATTAGTCAT CAACATTACA GTTTTCATGC TTCAGATATT TTACTGCTTG TGTCTTATT	1140
	GTGAGACAGC TTAAACAGA GTTGATGGTA CTTCAAATAT AGCTCATTGA TACTTAAGGG	1200
40	CANCTTCCTT GGGATGTGGG CTTTTTGAA GGAAAAAAT TNCCCCAAG GCAAATCCCA	1260
	GT	1262

45

(2) INFORMATION FOR SEQ ID NO: 261:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1179 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

50

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 261:

55

	GGCAAACCTT CCCCCAANGC TTCGAACTT GCAAGCCGAA ACCTTGAATC GTTAAAAGTT	60
	GGGTTGCGNC GGCGCCCTGG CCCGAAGAAG CGCAATTGGC GTTCCGCGAA CGTTGGCCCT	120
60	CAACGGCTCG GCAGCCAGCC ATGTCCTGCA CCCAGGACAG CGGCCCTGGG CTACAAGGAC	180

CTGGACCTCA TCTTCCTGCG CCGACCTGCG CGGGAAGGG GAGTTTCAGA CTGTGAAGGA 240
CGTCGTGCTG GACTGCCTGT TGGACTTCTT ACCCGAGGGG GTGAACAAAG AGAAGATCAC 300
5 ACCACTCACG CTCAAGGAAG CTTATGTGCA GAAAATGGTT AAAGTGTGCA ATGACTCTGA 360
CCGATGGAGT CTTATATCCC TGTCAAACAA CAGTGGCAAA AATGTGGAAC TGAAATTTGT 420
10 GGATTCCCTC CGGAGGCAGT TTGAATTCAG TGTAGATTCT TTTCAAATCA AATTAGACTC 480
TCTTCTGCTC TTTTATGAAT GTTCAGAGAA CCCAATGACT GAGACATTTC ACCCCACAAT 540
AATCGGGGAG AGCGTCTATG GCGATTTCGA GGAAGCCTTT GATCACCTTT GTAACAAGAT 600
15 CATTGCCACC AGGAACCCAG AGGAAATCCG AGGGGGAGGC CTGCTTAAGT ACTGCAACCT 660
CTTGGTGAGG GGCTTTAGGC CCGCCTCTGA TGAAATCAAG ACCCTTCAAA GGTATATGTG 720
20 TTCCAGGTTT TTCATCGACT TCTCAGACAT TGGAGAGCAG CAGAGAAAAC TGGAGTCCTA 780
TTTGCAAGC CACTTTGTGG GATTGGAAGA CCGCAAGTAT GAGTATCTCA TGACCCTTCA 840
TGGAGTGGTA AATGAGAGCA CAGTGTGCCT GATGGGACAT GAAAGAAGAC AGACTTTAAA 900
25 CCTTATCACC ATGCTGGCTA TCCGGGTGTT AGCTGACCAA AATGTCAATC CTAATGTGGC 960
TAATGTCACT TGCTATTACC AGCCAGCCCC CTATGTAGCA GATGCCAACT TTAGCAATTA 1020
30 CTACATGCA CAGGTTCAGC CAGTATTCAC GTGCCAGCAA CAGACCTACT CCACTTGGCT 1080
ACCTTCAAT TAAGAATCAT TAAAAATGT CCTGTGGGA AGCCATTTC AACAAGACAG 1140
GAGAGAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAGAGC 1179
35

(2) INFORMATION FOR SEQ ID NO: 262:

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1162 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

45

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 262:

GGCAAACTTT CCCCAANGC TTCGAACTT GCAAGCCGAA ACCTTGAATC GTTAAAAGTT 60
50 GGGTTGCGNC GCGCCCTGG CCCGAAGAAG CGCAATGGC GTTCCGCGAA CGTTGGCCCT 120
CAACGGCTCG GCAGCCAGCC ATGTCCTGCA CCCAGGACAG CGGCCCTGGG CTACAAGGAC 180
55 CTGGACCTCA TCTTCCTGCG CCGACCTGCG CGGGAAGGG GAGTTTCAGA CTGTGAAGGA 240
CGTCGTGCTG GACTGCCTGT TGGACTTCTT ACCCGAGGGG GTGAACAAAG AGAAGATCAC 300
ACCACTCACG CTCAAGGAAG CTTATGTGCA GAAAATGGTT AAAGTGTGCA ATGACTCTGA 360
60

CCGATGGAGT CTTATATCCC TGTCAAACAA CAGTGGCAAA AATGTGGAAC TGAAATTTGT 420
 GGATTCCTC CGGAGGCAGT TTGAATTCAG TGTAGATCT TTTCAAATCA AATTAGACTC 480
 5 TCTTCTGCTC TTTTATGAAT GTTCAGAGAA CCCAATGACT GAGACATTTC ACCCCACAAT 540
 AATCGGGGAG AGCGTCTATG GCGATTTCGA GGAAGCCTTT GATCACCTTT GTAACAAGAT 600
 10 CATTGCCACC AGGAACCCAG AGGAAATCCG AGGGGGAGGC CTGCTTAAGT ACTGCAACCT 660
 CTTGGTGAGG GGCTTTAGGC CCGCCTCTGA TGAAATCAAG ACCCTTCAAA GGTATATGTG 720
 TTCCAGGTTT TTCATCGACT TCTCAGACAT TGGAGAGCAG CAGAGAAAAC TGGAGTCCTA 780
 15 TTTGCAGAAC CACTTTGTGG GATTGGAAGA CCGCAAGTAT GAGTATCTCA TGACCCTTCA 840
 TGGAGTGGTA AATGAGAGCA CAGTGTGCCT GATGGGACAT GAAAGAAGAC AGACTTTAAA 900
 CCTTATCACC ATGCTGGCTA TCCGGGTGTT AGCTGACCAA AATGTCATT CTAATGTGGC 960
 20 TAATGTCACT TGCTATTACC AGCCAGCCCC CTATGTAGCA GATGCCAACT TTAGCAATTA 1020
 CTACATTGCA CAGGTTGAGC CAGTATTCAC GTGCCAGCAA CAGACCTACT CCACTTGGCT 1080
 25 ACCCTGCAAT TAAGAATCAT TTA AAAATGT CCTGTGGGGA AGCCATTTC GACAAGACAG 1140
 GAGAGAAAAA NAANGAAAAG AG 1162

30

(2) INFORMATION FOR SEQ ID NO: 263:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 735 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 263:

CGGGCTGGGT ATTTGCCTCG CACCATGGCG CCCAAGGGCA AAGTGGGCAC GAGAGGGAAG 60
 AAGCAGATAT TTGAAGAGAA CAGAGAGACT CTGAAGTTCT ACCTGCGGAT CATACTGGGG 120
 45 GCCAATGCCA TTTACTGCCT TGTGACGTTG GTCTTCTTTT ACTCATCTGC CTCATTTTGG 180
 GCCTGGTTGG CCTTGGGCTT TAGTCTGGCA GTGTATGGGG CCAGCTACCA CTCTATGAGC 240
 50 TCGATGGCAC GAGCAGCGTT CTTCTGAGGA TGGGGCCCTG ATGGATGGTG GCACGAGCTC 300
 AACATGGAGC AGGGCATGGC AGAGCACCTT AAGGATGTGA TCCTACTGAC AGCCATCGTG 360
 CAGGTGCTCA GCTGCTTCTC TCTCTATGTC TGGTCCCTTCT GGCTTCTGGC TCCAGGCCGG 420
 55 GCCCTTTACC TCCTGTGGGT GAATGTGCTG GGCCCCGGT TCACTGCAGA CAGTGGCACC 480
 CCAGCACCAG AGCACAATGA GAAACGGCAG CGCCGACAGG AGCGCGGCA GATGAAGCGG 540
 60 TTATAGCCAT TGACATTGTG GCCACAGGCC ACTGGCCCTG GGTGGCTCTG TCAGGGTGCA 600

5 CAGCCCCTCA TGCCTGGAGC AATGAGGGTC TAGTCCAGGG GCCAAAAGCA GTCTGAGGTA 660
TTGGGTATAC TTATACTCTA TAGGGTCGTT GAATAAATGG CTTAGAATGT GAAAAAAAAA 720
AAAAAAAAA ATTTT 735

10

(2) INFORMATION FOR SEQ ID NO: 264:

15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 783 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 264:

AAGTGCATGA GCTGCCGATG TGGTGCTTAG TGATTGCCGT TTCGGTCGCT CTCCCGTGTT 60
TCCCGGGCTG GGTATTTGCC TCGCACCATG GCGCCCAAGG GCAAAGTGG CACGAGAGGG 120
25 AAGAAGCAGA TATTTGAAGA GAACAGAGAG ACTCTGAAGT TCTACCTGCG GATCATACTG 180
GGGGCCAATG CCATTTACTG CCTGTGACG TTGGTCTTCT TTTACTCATC TGCCTCATTT 240
TGGCCTGGT TGGCCTGGC TTTAGTCTGG CAGTGTATGG GGCCAGCTAC CACTCTATGA 300
30 GCTCGATGGC ACGAGCAGCG TTCTCTGAGG ATGGGGCCCT GATGGATGGT GGCATGGACC 360
TCAACATGGA GCAGGGCATG GCAGAGTGAG TGTCCCCAC CGCCAGCCCA GGCACCTTAA 420
35 GGATGTGATC CTACTGACAG CCATCGTGCA GGTGCTCAGC TGCTTCTCTC TCTATGCTCG 480
GTCCTTCTGG CTCTCTGGCTC CAGGCCGGGC CCTTTACCTC CTGTGGGTGA ATGTGCTGGG 540
CCCCTGGTTC ACTGCAGACA GTGGCACCCC AGCACCAGAG CACAATGAGA AACGGCAGCG 600
40 CCGACAGGAG CGGCGGCAGA TGAAGCGGTT ATAGCCATTG ACGATTTRGC SACNRGCCAC 660
TGGCCCTGGG TGGCTCTGTC AGGGTGACA GCCCCTCATG CCTGGAGCAA TGAGGGTCTA 720
45 GTCCAGGGGC CAAAAGCAGT CTGAGGTATT GGGTATACTT ATACTCTATA GGGTCGTTGA 780
ATA 783

50

(2) INFORMATION FOR SEQ ID NO: 265:

55 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1638 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 265:

	GGCACGAGGC GCGGCAGCG GTGGCGGCGG CGCCCCCGG CGGGAGCCGT NCCCTTTCCC	60
5	GTGGGGGAGC GCGGGGYCGG GGYCCAGGG ANCCCGGMC ACGGAGAGCG GGAAGAGGAT	120
	GGATTGCCCG GCCCTCCCC CCGATGGAA GAAGGAGGAA GTGATCCGAA AATCTGGGCT	180
	AAGTGCTGGC AAGAGCGATG TCTACTACTT CAGTCCAAGT GGTAAAGAAGT TCAGAAGCAA	240
10	GCCTCAGTTG GCAAGGTACC TGGGAAATAC TGTGTATCTC AGCAGTTTGT ACTTCAGAAC	300
	TGGAAAGATG ATGCCTAGTA AATTACAGAA GAACAAACAG AGACTGCGAA ACGATCCTCT	360
15	CAATCAAAAT AAGGGTAAAC CAGACTTGAA TACAACATTG CCAATTAGAC AAACAGCATC	420
	AATTTTCAAA CAACCGGTAA CCAAAGTCAC AAATCATCCT AGTAATAAAG TGAAATCAGA	480
	CCCACAACGA ATGAATGAAC AGCCACGTCA GCTTTTCTGG GAGAAGAGGC TACAAGGACT	540
20	TAGTGCATCA GATGTAACAG AACAAATTAT AAAAACCATG GAACTACCCA AAGGTCTTCA	600
	AGGAGTTGGT CCAGGTAGCA ATGATGAGAC CCTTTTATCT GCTGTTGCCA GTGCTTTGCA	660
25	CACAAGCTCT GCGCCAATCA CAGGGCAAGT CTCGCTGCT GTGGAAAAGA ACCCTGCTGT	720
	TTGGCTTAAC ACATCTCAAC CCCTCTGCAA AGCTTTTATT GTACAGATG AAGACATCAG	780
	GAAACAGGAA GAGCGAGTAC AGCAAGTACG CAAGAAATTG GAAGAAGCAC TGATGGCAGA	840
30	CATCTTGTCTG CGAGCTGCTG ATACAGAAGA GATGGATATT GAAATGGACA GTGGAGATGA	900
	AGCCTAAGAA TATGATCAGG TAACTTTCGA CCGACTTTCC CCAAGAGAAA ATTCTTAGAA	960
35	ATTGAACAAA AATGTTTCCA CTGGCTTTTG CCTGTAAGAA AAAAAATGTA CCCGAGCACA	1020
	TAGAGCTTTT TAATAGCACT AACCAATGCC TTTTATAGTG TATTTTGTAT GTATATATCT	1080
	ATTATTCAAA AAATCATGTT TATTTTGAGT CCTAGGACTT AAAATTAGTC TTTTGTAAATA	1140
40	TCAAGCAGGA CCCTAAGATG AAGCTGAGCT TTTGATGCCA GGTGCAATCT ACTGGAAATG	1200
	TAGCACTTAC GTAAACATT TGTTCCCCC ACAGTTTAA TAAGAACAGA TCAGGAATTC	1260
45	TAAATAAAAT TCCCAGTTAA AGATTATTGT GACTTCACTG TATATAAACA TATTTTATA	1320
	CTTTATTGAA AGGGGACACC TGTACATTCT TCCATCCTCA CTGTAAAGAC AAATAAATGA	1380
	TTATATTAC AGACTGATTG GAATTCCTTC TGTGAAAAG CACACACAAT AAAGAACCCC	1440
50	TCGTTAGCCT TCCTCTGATT TACATTCAAC TCTGATCCCG GGGCCTTAGG TTTGACATGG	1500
	GAGGTGGGAG GAAGATAGCG CATATATTGT CAGTATGAAC TATTGCCTCT GGGACGTTGT	1560
55	GAGGAATTGT GCTTTCACCA GAATTTCTAA GGATTTCTGG CTTAAATATC ACCTAGCCTG	1620
	TGGAATTTT TTTTCCCT	1638
60		

(2) INFORMATION FOR SEQ ID NO: 266:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 1455 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 266:

10 CGTGCCTACT GCCATGCAGG TACCGGGTCC GGAATTCCTCA GGGTCGACCC ACGCGTCCGC 60
TCAGTTGGCA AGGTACCTGG GAAATACTGT TGATCTCAGC AGTTTGTACT TCAGAACTGG 120
15 AAAGATGATG CCTAGTAAAT TACAGAAGAA CAAACAGAGA CTGCGAAACG ATCCTCTCAA 180
TCAAAATAAG GGTAAACCAG ACTTGAATAC AACATTGCCA ATTAGACAAA CAGCATCAAT 240
TTTCAAAACA CCGGTAACCA AAGTCACAAA TCATCTAGT AATAAAGTGA AATCAGACCC 300
20 ACAACGAATG AATGAACAGC CACGTCAGCT TTTCTGGGAG AAGAGGCTAC AAGGACTTAG 360
TGCATCAGAT GTAACAGAAC AAATTATAAA AACCATGGAA CTACCCAAAG GTCTTCAAGG 420
25 AGTTGGTCCA GGTAGCAATG ATGAGACCTT TTTATCTGCT GTTGCCAGTG CTTTGACACAC 480
AAGCTCTGCG CCAATCACAG GGCAAGTCTC CGCTGCTGTG GAAAAGAACC CTGCTGTTTG 540
GCTTAACACA TCTCAACCCC TCTGCAAAGC TTTTATTGTC ACAGATGAAG ACATCAGGAA 600
30 ACAGGAAGAG CGAGTACAGC AAGTACGCAA GAAATTGGAA GAAGCACTGA TGGCAGACAT 660
CTTGTGCGCA GCTGCTGATA CAGAAGAGAT GGATATTGAA ATGGACAGTG GAGATGAAGC 720
35 CTAAGAATAT GATCAGGTAA CTTTCGACCG ACTTTCCTCA AGAGAAAATT CCTAGAAATT 780
GAACAAAAT GTTCCACTG GCTTTTGCCT GTAAGAAAAA AAATGTACCC GAGCACATAG 840
AGCTTTTAA TAGCACTAAC CAATGCCTTT TTAGATGTAT TTTTGATGTA TATATCTATT 900
40 ATTCAAAAA TCATGTTTAT TTTGAGTCCT AGGACTTAAA ATTAGTCTTT TGTAAATATCA 960
AGCAGGACCC TAAGATGAAG CTGAGCTTTT GATGCCAGGT GCAATCTACT GGAAATGTAG 1020
45 CACTTACGTA AAACATTTGT TTCCCCACA GTTTTAATAA GAACAGATCA GGAATTCTAA 1080
ATAAATTTC CAGTTAAAGA TTATTGTGAC TTCCTGTAT ATAAACATAT TTTTATACTT 1140
TATTGAAAGG GGACACCTGT ACATTCTTCC ATCCTCACTG TAAAGACAAA TAAATGATTA 1200
50 TATTACAGA CTGATTGGAA TTCTTCTGT TGAAAAGCAC ACACAATAAA GAACCCCTCG 1260
TTAGCCTTCC TCTGATTTAC ATTCAACTCT GATCCCGGG CCTTAGGTTT GACATGGGAG 1320
55 GTGGGAGGAA GATAGCGCAT ATATTTCAG TATGAACTAT TGCTCTGGG ACGTTGTGAG 1380
GAATTGTGCT TTCACCAGAA TTCTAAGGA TTCTGGCTT AAATATCACC TAGCCTGTGG 1440
60 TAATTTTTTT TCCCT 1455

(2) INFORMATION FOR SEQ ID NO: 267:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1086 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

10

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 267:

15	CGCCTGCAGT ACCGGTCCGG AATTCCTGGG TCGACCCACG CGTCGCTGAC CCAGGAGAAG	60
	CTGCCTGTCT ACATCAGCCT GGGCTGCAGC GCGCTGCCGC CGCGGGGCGG GCAGCTGAAC	120
	TATGTGCTCT TCAGGGCGGG CACCGTGTG CAITCATCTT TGTACCCCA GCATCTAGCA	180
20	GTGTGGCAT GTAGTAGGCA CTCAAGAAAT GTGTGTTGAA TGAACGATGC CTGTGACAAG	240
	CAAGCGGACT TTATTCTTTC CTGACCCTTG CTCCTATGAC ACACCTCCTC CTGACTGCCA	300
25	CTGTCACTCC TTCAGAGCAG AACTCCTCTA GGGAACTGG ATGGGAAACA GCCATGGCCA	360
	AGGACATCCT GGGTGAAGCA GGGCTACACT TTGATGAACT GAACAAGCTG AGGGTGTGG	420
	ACCCAGAGGT TACCCAGCAG ACCATAGAGC TGAAGGAAGA GTGCAAAGAC TTTGTGGACA	480
30	AAATGGCCA GTTTCAGAAA ATAGTTGGTG GTTTAATTGA GCTTGTGAT CAACTTGCAA	540
	AAGAAGCAGA AAATGAAAAG ATGAAGGCCA TCGGTGCTCG GAACTTGCTC AAATCTATAG	600
35	CAAAGCAGAG AGAAGCTCAA CAGCAGCAAC TTCAAGCCCT AATAGCAGAA AAGAAAATGC	660
	AGCTAGAAAG GTATCGGGTT GAATATGAAG CTTTGTGTAA AGTAGAAGCA GAACAAAATG	720
	AATTTATTGA CCAATTTATT TTTCAGAAAT GAACTGAAAA TTTGCTTTT ATAGTAGGAA	780
40	GGCAAAACAA AAAAAAGCCT CTCAAAACCA AAAAAACCTC TGTAGCATTC CAGCGGCTTG	840
	ACCAATGACC TATGTCACAA GAGGTGGCGT GTAAGGAATG CAGCCCCCTG AAGACAGCAC	900
45	TACAAGTCTG GGGAGCCAG TTTTAACATC AGTGCACAGC TGCTGCTGGT GGCCCTGCAG	960
	TGTACGTTCT CACCTCTTAT GCTTAGTTGG AACTAAGCAG TTTGTAACT TTCATCCTTT	1020
	TTTTTGTAAT TTCACAAAGC TTTGGAAGGA GARGCAATAA ATTTTGTGTT TCNAAATGGC	1080
50	TTGATG	1086

55 (2) INFORMATION FOR SEQ ID NO: 268:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1003 base pairs

(B) TYPE: nucleic acid

60

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 268:

5 GGCACGGGAG CAGCCGGGCT GGTCTGCTG CGAGCCGGCG GCCCGGAGTG GGGCGGCGGA 60
 GCAAACATGA ACCTTGGAGT TGCCACAGT GAAGTGAATC CAAATACCCG TGTCAATGAAC 120
 AGCCGGGGTA TGTGGCTGAC ATATGCATTG GGAGTGGCT TGCTTCATAT TGTCTTACTC 180
 10 AGCATTCCTT TCTTCAGTGT TCCTGTTGCT TGGACTTTAA CAAATATTAT ACATAATCTG 240
 GGGATGTACG TATTTTTCGA TGCAGTGAAA GGAACACCTT TCGAAACTCC TGACCAGGGT 300
 15 AAAAGCAAGG CTCCTAACTC ATTGGGAACA ACTGGACTAT GGAGTACAGT TTACATCTTC 360
 ACGGAAGTTT TTCACAATTT CTCCAATAAT TCTATATTTT CTGGCAAGTT TCTATACGAA 420
 GTATGATCCA ACTCACTTCA TCCTAAACAC AGCTTCTCTC CTGAGTGTAC TAATTCCTCA 480
 20 AATGCCACAA CTACATGGTG TTCGGATCTT TGGAAATTAAT AAGTATTGAA ATGTTTTCGAA 540
 ACTGAAAAAA AATTTTACAG CTACTGAATT TCTTATAAGG AAGGAGTGGT TAGTAACTG 600
 25 CACTGTTTCT CTGATAATGT GAAATGAGAA GTATTTACAT TGGAGGGCCA ATGGCTGGTC 660
 CTTCAAGTGC TGTTTTCGAG TGCAGATTTT CATTAATGA TGCCTCTGTT TAATACACCT 720
 GGTACATTTT TGAAGAGGGG CTTTATAAGC AGGCTGGGCA GCGCCAGCTT ATAAGTTAAA 780
 30 GGGCATCACA GTGAGGTGT AGTAGATAAA TTCAAGGAAA TAAGAGATTT GTAAGAACT 840
 AGGACCAGCT TAACTTATAA TGAATGGCA TTGTGTTAAG AAAAGAACAT TTCCAGTCAT 900
 35 TCAGCTGTGG TTATTTAAAG CAGACTTACA TGTAACCCG AATCCTCTCT ATACAAGTTT 960
 ATTAAAGATT ATTTTATTA CCGTAAAAA AAAAAAAAA AAA 1003

40

(2) INFORMATION FOR SEQ ID NO: 269:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 1234 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 269:

ATCAGCATCT ACAAGTAGCA TATTTGGAT GGTGTTGTG TGCTACTTCA AAGTAACTAG 60
 GAAAAATAA TCCTCGAAC ACAGGTACCT TGTCAATGCA GAATGGGGG TGTAGGTTG 120
 55 CCAGTGTAT CAGTGTGAT TCATTTTATT ACTTCCTACA GAGCAAACAT GAACGTTGGA 180
 GTTGCCACA GTGAAGTGAA TCCAAATACC CGTGTATGA ACAGCCGGG TATGTGGCTG 240
 60 ACATATGCAT TGGGAGTTGG CTGCTTCAT ATTGTCTTAC TCAGCATTC CTCTTCAGT 300

GTTCCTGTTG CTGGAAGTTT AACAAATATT ATACATAATC TGGGGATGTA CGTATTTTGT 360
 CATGCAGTGA AAGGAACACC TTTCGAAACT CCTGACCAGG GTAAAGCAAG GCTCCTAACT 420
 5 CATTGGGAAC AACTGGACTA TGGAGTACAG TTTACATCTT CACGGAAGTT TTTACAATT 480
 TCTCCAATAA TTCTATATTT TCTGGCAAGT TTCTATACGA AGTATGATCC AACTCACTTC 540
 10 ATCCTAAACA CAGCTTCTCT CCTGAGTGTA CTAATTCCCA AAATGCCACA ACTACATGGT 600
 GTTCGGATCT TTGGAATTAA TAAGTATTGA AATGTTTGA AACTGAAAAA AAATTTTACA 660
 GCTACTGAAT TTCTTATAAG GAAGGAGTGG TTAGTAAACT GCACTGTTTC TSTGATAATG 720
 15 TGAAATGAGA AGTATTTACA TTGGAGGGCC AATGGCTGGT CCTTCAAGTG CTGTTTTGAA 780
 GTGCAGATTT CCATTAAATG ATGCCTCTGT TTAATACACC TGGTACATTT CTGAAGAGGG 840
 20 GCTTTATAAG CARGCTGGGC AGGCCAGCT TATAAGTTAA AGGGCATCAC AGTGAGGGTG 900
 TAGTAGATAA ATTCAAGGAA ATAAGAGATT TGTAAGAAAC TAGGACCAGC TTAACCTATA 960
 ATGAATGGGC ATTGTGTTAA GAAAAGAACA TTCCAGTCA TTCAGCTGTG GTTATTTAAA 1020
 25 GCAGACTTAC ATGTAAACCG GAATCCTCTC TATACAAGTT TATTAAAGAT TATTTTATT 1080
 ACCRTACATA TTTCKCTGT TTTATGTAAG YGGATGTATA TCCTCTGTGTT TTATACAAGC 1140
 30 CAGTCCCAC TTATGAGGGT ACTTTTTTGG TTTTGCTGGG CTTAATATTG TGTATTGGTC 1200
 AATGAGGCCA TTTTACANT TATTAACGTT ACAG 1234

35

(2) INFORMATION FOR SEQ ID NO: 270:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 574 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 270:

NGAGGTGCGT TCTGAGCCGT CTGTCCTGCG CCAAGATGCT TCAAAGTATT ATTAAAAACA 60
 TATGGATCCC CATGAAGCCC TACTACACCA AAGTTTACCA GGAGATTTGG ATAGGAATGG 120
 50 GGCTGATGGG CTTTCATCGTT TATAAAATCC GGGCTGCTGA TAAAAGAAGT AAGGCTTTGA 180
 AAGCTTCAGC GCCTGCTCCT GGTCACTACT AACCAGATTT ACTTGGAGTA CATGTGAAAG 240
 55 AAAACGTCAG TCTGCCTGTA AATTTCAGCA AGCCGTGTTA GATGGGGAGC GTGGAACGTC 300
 ACTGTACACT TGTATAAGTA CCGTTTACTT CATGGCATGA ATAAATGGAT CTGTGAGATG 360
 CACTGCTACC TGGTACTGCT TTCAGTGTGT TCCCCCTCAG CCCTCCGGCG TGTACGGCAT 420
 60

ACTCTGAGTA GATAATTGT CATGCAGCGC ATGCAATCAG AATCTCACTG AGCCACCCAT 480
CATTGTGAAA TAATTACCTC AGTTGTACAG GACTTGGTGA TCAGGATCCA GGCACCTACT 540
5 TGTATTCTAC TGCTCAATAA ACGTTTATTA AACT 574

10 (2) INFORMATION FOR SEQ ID NO: 271:

(i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 1731 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 271:

20 GCTGCAAGGT GCGCCTCGTG CCGCTGCAGA TCCAGCTCAC TACCTTGGGA AATCTTACAC 60
CTTCAAGCAC TGTGTTTTTC TGCTGTGATA TGCAGGAAAG GTTCAGACCA GCCATCAAGT 120
ATTTTGGGGA TATTATTAGC GTGGGACAGA GATTGTTGCA AGGGGCCCGG ATTTTAGGAA 180
25 TTCTTGTTAT TGTAACAGAA CAATACCCTA AAGGTCTTGG GAGCACGGTT CAAGAAATTG 240
ATTTAACAGG TGTAAGAACTG GTACTTCCAA AGACCAAGTT TTCAATGGTA TTACCAGAAG 300
30 TAGAAGCGGC ATTAGCAGAG ATTCCCGGAG TCAGGAGTGT TGTATTATTT GGAGTAGAAA 360
CTCATGTGTG CATCCAACAA ACTGCCCTGG AGCTAGTTGG CCGAGGAGTC GAGGTTTACA 420
TTGTTGCTGA TGCCACCTCA TCAAGAAGCA TGATGGACAG GATGTTTGCC CTCGAGCGTC 480
35 TCGCTCRARC CNGGGATCAT AGTGACCACG AGTGNAGGCT GTTCTGCTTC AGCTGGTAGC 540
TGATAAGGAC CATCCAAAT TCAAGGAAAT TCAGAATCTA ATTAAGGCGA GTGCTCCAGA 600
40 GTCGGGTCTG CTTTCCAAAG TATAGGACAT TTGAAGAACT GGTATGCTAC TCACTGGTGA 660
AGGACAGTCA GGTGAAGGAC TGTAAGCCCA CACAAGCTCT TCTTATCTCT ACTAGAATTA 720
AAATGTTAAG TCAAAAACGG CTCCTTTTTT GCGCCTCCTA GTGAACITAA CCAGCTAGAC 780
45 CATTTGAGTA CCAGCATTTA GTTACAAACG TCAAAGGCTT CCGGTGCTGC TTACCTTCCT 840
TTTTTGTTAA TGTGCTTTTA TTTATTAAAA AAAATTACAA TGAAGATGCC TGTTTTGTCT 900
50 CTACTGTGTA CTCTGATCGT ATCTTTCCAA AGTGCAGACT CTTGTGAAGT TTTCTTAAAT 960
TGTTCACTTT AAAGAAAATG ACGTACCAAC AATGATTGG CTTTATATT ACTGTAAGAT 1020
GTTATAATGT TAATGTGGAT GTAGTGCTTT TACTTTACAG ATTGATTGGA ATAAGATTAT 1080
55 TGCATATGAA TTTACCCACA GGACTCTGAA TCAITTTACC CACTCCCTC ACAATGTTGT 1140
CCACTTAGTG AGTTGCATTG ATCTATCCGT ACCAAATGAT GTTGAATAAT TACATATCTT 1200
60 TCTKGACTAT ACTGATTTCT TATTTTGGTC ACTATTACTA AATCTCTGTT AATATTCTCT 1260

CTTTTAACTG AAAAGGGATG GGATAGAAGG GTTTGCAATG CCATATTATT GGTGGAGGGC 1320
TGTTTTAAACA TCTTTGAAGT ATGGCTTGCT GAATATCTTT ACCAACATCT TGAATATATA 1380
5 TTCTAGTGTC CACAAGATTT AGCAAAAAGA TAAAGCTTGG GTGGAATATC ATTTTAAAAAT 1440
GTTCATGTTT TGTTCTATAT TTTCTTCACC TACTCTCCAA ATATTGTAAT GCAAAAAGTC 1500
10 TCAGTAATGA TTTGGTAGTA TTAATTTTGT GGTCATTTGT TCTCTTCGAT AAATTTATTT 1560
TCATTAAATA CTTRTTAGAG GGTTTGAAGT TGTTTTCCTT ATATGTGAAA TGTGAAACTG 1620
CTGTCTTTTA TATTAAAGTA ATTAAAGAAA ATGTATTGTG ATTGAAATTA TTTTGNCCTC 1680
15 CACAAGATGG CTCTATGAGT ATTCTTCCAG GGATTCTAAT ATTTATTTAA G 1731

20

(2) INFORMATION FOR SEQ ID NO: 272:

(i) SEQUENCE CHARACTERISTICS:

25

- (A) LENGTH: 1320 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 272:

CTGCTTAGGA AGAGAAGGTC AGAGTTCGCG GGGGCAGAGG CATTCCTGCC GCTGGCCCAG 60
TCACTATGTA GTGGAGGGGC AGACACCCTC CCGCAAATTC TGGAAGGTTC TTAGTCTCGA 120
35 CTAGGGCAGT AGCCCAGGAC TCCTAGTCGC CGGCTTCAGG TCACTGCCGG CTGAACGGAG 180
CTGCCGTCGC CATGTTTGGC TGCTTGGTGG CGGGGAGGCT GGTGCAAACA GCTGCACAGC 240
AAGTGGCAGA GGATAAATTT GTTTTGGACT TACCTGATTA TGAAAGTATC AACCATGTTG 300
40 TGGTTTTAT GCTGGGAACA ATCCCATTTT CTGAGGGAAT GGGAGGATCT GTCTACTTTT 360
CTTATCCTGA TTCAAATGGA ATGCCAGTAT GGMAACTCCT AGGATTGTGC ACGAATGGGA 420
45 AGCCAAGTGC CATCTTCAAA ATTTCAAGTC TTAAATCTGG AGAAGGAAGC CAACATCCTT 480
TTGGAGCCAT GAATATTGTC CGAACTCCAT CTGTTGCTCA GATTGGAATT TCAGTGGAAAT 540
TATTAGACAG TATGGCTCAG CAGACTCCTG TAGGTAATGC TGCTGTATCC TCAGTTGACT 600
50 CATTCACTCA GTTCACACAA AAGATGTTGG ACAATTTCTA CAATTTTGCT TCATCAITTG 660
CTGTCTCTCA GGCCCAGATG ACACCAAGCC CATCTGAAAT GTTCATTCCG GCAAATGTGG 720
55 TTCTGCAAAT GGTATGAGGC ATNITCTGTC TCCAATATTA AGGCTTTTTA TAACTGAATA 780
TCTATTTTGT CTATGAATAT ATTCTTTTTT TGACATTTAA ACATATTCTT TTATTGTGAA 840
CATCAGCACT GCATGCCATT AAAGTATGTA CTATAGAGAT CTGATGAGAA ACAGTTCTTA 900
60

CCCTAAATAT TTTGTTATAT TGTCGCCATT ATGAATTTAT AAAGACAGGA AAATATAGTT 960
GCCTATGTTT TAGGGACCAC TATTAAAGCT TATAAATATT TGTGTATTTT CATTTAGAAG 1020
5 TACCATCTAT GAGAGTAGTT TATACTGCAC TGTGTACATG AATGGCTAAT GAATCTATTT 1080
TCCAACITTC CCGTGTITTA TAGATATTC TTTTCACTTT GAGTATCCTA GAGATGGGAG 1140
GATGCCCTAGG AAGAGTTTGT TGAGAAGTGG TACCATGGTG TAGCATGGGA GAGCATTGGG 1200
10 AATGCACTAG GTTTGAATTT GGCATAATGG TAGCTATGTG ACCCTGAGCA AATTTCTCTC 1260
ATCTGCTCAT CTGANGAATG AGGAAATAGG AGTGAATTTG ATNTTTCCTA GGTCCNTCTA 1320
15

(2) INFORMATION FOR SEQ ID NO: 273:

20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 515 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 273:

CCCTGGAGAG GGGCTGCTGT GCCAGCTTGG GGAGGGTCTG GGATGGGGCT GCCCCTGATG 60
30 GCCCTGATGT GGAGTACCTT GCCAGCATCT GCTGGGGTGA ACTTTATTTT AGCCCTTCCC 120
TTGTTGYTCT TATGAAGAAC AGAGGAGGGG TGGGCAGGTC AGTGATGTCA GCAGTGAGTA 180
TTCCCAGCAC AGCGGCTCTG GAAGAGGCAT GAGGCATTTT TTTCAGGAAA TGRTCATTTAT 240
35 TCAGCCAGAA GGCAITTCATT AAGTAAGTCC TGACTTTGTG CCCAGCTCTG TGTATATAGGC 300
CCTTGGCGAG ACTCAGGAGG GGCARAGGAC GCTAGKTTKT AGWTAACACG GAACCTCARA 360
40 GGWTATATGG TCCAAGAAGA CCCGGGGGCG GTGAAAACCC TGTGGACTAA TGCTCACGGG 420
AGCCCCGAGGT CACACTTTGA CTTTGCTACC ATGGGCTGTG TCTANGNACG TATATATGCT 480
45 GCGTAATTAT TACAGAGGCA GTCCATGTGC ATTGT 515

(2) INFORMATION FOR SEQ ID NO: 274:

50 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2995 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
55 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 274:

60 TGACACCCAT AAGGAATTCA TGAAGAAAGT AGAAGAAAAG CGAGTGGACG TTAACCTCAGC 60

	AGTAGCCATG GGAGAAGTCA TCCTGGCTGT CTGCCACCCC GATTGCATCA CAACCATCAA	120
	ACACTGGATC ACCATCATCC GAGCTCGCTT CGAGGAGGTC CTGACATGGG CTAAGCAGCA	180
5	CCAGCAGCGT CTTGAAACGG CCTTGTGAGA ACTGGTGGCT AATGCTGAGC TCCTGGAAGA	240
	ACTTCTGGCA TGGATCCAGT GGGCTGAGAC CACCCTCATT CAGCGGGATC AGGAGCCAAAT	300
10	CCCGCAGAAC ATTGACCGAG TTAAAGCCCT TATCGCTGAG CATCAGACAT TTATGGAGGA	360
	GATGACTCGC AAACAGCCTG ACGTGGACCG GGTCAACCAAG ACATACAAAA GGAAAAACAT	420
	AGAGCCTACT CACGCGCCTT TCATAGAGAA ATCCCGCAGC GGAGGCAGGA AATCCCTAAG	480
15	TCAGCCAACC CCTCCTCCCA TGCCAATCCT TTCACAGTCT GAAGCAAAAA ACCCACGGAT	540
	CAACCAGCTT TCTGCCCGCT GGCAGCAGGT GTGGCTGTTA GCACTGGAGC GGCAAAGGAA	600
20	ACTGAATGAT GCCTTGGATC GGCTGGAGGA GTTGAAAGAA TTTGCCAACT TTGACTTTGA	660
	TGTCTGGAGG AAAAAGTATA TCGTTGGAT GAATCACAAA AAGTCTCGAG TGATGGATTT	720
	CTTCCGGCGC ATTGATAAGG ACCAGGATGG GAAGATAACA CGTCAGGAGT TTATCGATGG	780
25	CATTTTAGCA TCCAAGTTCC CCACCACCAA GTTAGAGATG ACTGCTGTGG CTGACATTTT	840
	CGACCGAGAT GGGGATGGTT ACATTGATTA TTATGAATTT GTGGCTGCTC TTCATCCCAA	900
30	CAAGGATGCG TATCGACCAA CAACCGATGC AGATAAAATC GAAGATGAGG TTACAAGACA	960
	AGTGGCTCAG TGCAAATGTG CAAAAGGTT TCAGGTGGAG CAGATCGGAG AGAATAAATA	1020
	C CGGTCTCTC CTCGGCAATC AGTTTGGGGA TTCTCAGCAG TTGCGGCTGG TCCGTATTCT	1080
35	GCGCAACCGT GATGGTTCGC GTTGGTGGAG GATGGATGGC CTTGGATGAA TTTTTAGTGA	1140
	AAAATGATCC CTGCCGAGCA CGAGGTAGAA CTAACATGA ACTTAGAGAG AAATTCATCC	1200
40	TACCAGAGGG AGCATCCCAG GGAATGACCC CCTTCCGCTC ACGGGGTCGA AGGTCCAAAC	1260
	CATCTTCCCC GGCAGCTTCC CCTACTCGTT CCAGCTCCAG TGCTAGTCAG AGTAACCACA	1320
	GCTGTACATC CATGCCATCT TCTCCAGCCA CCCAGCCAG TGAACCAAG GTTATCCCAT	1380
45	CATCAGGTAG CAAGTTGAAA CGACCAACAC CAACTTTTCA TTCTAGTCGG ACATCCCTTG	1440
	CTGGTGATAC CAGCAATTAG TTCTTCCCCG GCCTCCACAG GTGCCAAAAC TAATCGGGCA	1500
50	GACCCATAAA AGTCTGCCAG TCGCCCTGGG AGTCGGGCTG GGAGTCGAGC CGGGAGTCGA	1560
	GCCAGCAGCC GCGGAGGAAG TGACGCTTCT GACTTTGACC TCTTAGAGAC GCATTGCTTG	1620
	TTCCGCACT TCAGAAAGCA GCGCTGCAGG GGGCCAAGGC AACTCCAGGA GAGGGCTAAA	1680
55	CAAACCTTCC AAAATCCCAA CCATGTCTAA GAAGACCACC ACTGCCTCCC CCAGGACTCC	1740
	AGGTCCCAAG CGATAACACT GTCTAAGCAC CCCAAGCCA CTATCCACTT TGAATCCTGC	1800
60	TCCATACATT GGGTGTATAT TTATTCTGAA CGGGAGAAGT TATATTGTTA AAAGTGTAAG	1860

AGAATAATTG TGTATGAAG CTGCCTTATT TTTTTCCTTT TGTAAAGTTA CTATTTTCAT 1920
 GTGAATATTT ATGTAGATAA AATTTCCTTC CTGGTAACCC TGTAATGGAT GGGGCCCAGA 1980
 5 AATGAAATAT TTGAGAAAA CAAGTGAAAA GGTCAAGATA CAAATGTGTA TTAACAAAAA 2040
 AAAAGCCTAT TAATAGGGTT TCTGCGCGT GCAGGGTTGT AAACCTGCTT TATCTTTTAG 2100
 GATTATTCCT AAATGCATCT TCTTTATAAA CTTGACTTGC TATCTCAGCA AGATAAATTA 2160
 10 TATTAACAAA ATAAGAATCC TGCAGTGTTC AAGGAACCTT TTTTGTGTAA ATCACGGACA 2220
 CCTCAATTAG CAAGAACTGA GGGGAGGGCT TTTTCCATTG TTTAATGTTT TGTGATTTTT 2280
 15 AGCTAAAGAG AGGGAACCTC ATCTAAGTAA CATTTCACA TGGATACAGC AAAAGGAGTT 2340
 CATTGCAATA CTGCTTTTGG ATATTGTTTC AGTACTGGGT GTTTAAAGGA CAAATAGCTG 2400
 CTAGAATTCA GGGGTAAATG TAAGTGTTCA GAAAACGTCA GAACATTTGG GGTTTTAAAC 2460
 20 TGATTTGTTG CTCCCTATCC AGCCTAGACA CCAGTAACTC TTGTGTTTAC CAGGACCCAG 2520
 ACCCTTGCCA AGGGATAGGC TCGTTGGTGA CATGTGGAAT TTCAGATTTG TTTTATCCAC 2580
 25 TTTTGTGCT ATTTATTTAA ATGGTCGATC AACTTCCAC AAACGAGGA ATGAATTCCA 2640
 CGAGCCTGTT CTGAAAATGT GGACGTAAGA CAAACACGTG CTCGTCCTTT AATGGAGTTC 2700
 ACCAGCACAC TTGTTAACCA GTCCTGTTTG CTTCGTCCTT TTTTGTGCG TAATAAAGTC 2760
 30 AACTGACCAA GTGACCATGA AAAGGGGCTG TCTGGGGCTC CTGTTTMTTA GCTGCTGTTT 2820
 TTCAGCTCCG ACCATGTTGC TGTGTGATTA TCTCAATGG TTTTAATTGA GGCAGAAACT 2880
 35 GAAGCTCTAC CAATGAACTG TTTAGAAACA AGACACACTT TTGTATTAAA ATTGCTTGCA 2940
 GTAACAAAAA AAAAAAAAAA AAAAAAAAAA AAAAACTCG AGGGGGGCCC GGTAC 2995

40

(2) INFORMATION FOR SEQ ID NO: 275:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1990 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 275:

GGGACCCGCG CGSCTCCCGG GGATGGTGAG CAAGGCGCTG CTGCNWCCTG TCTGCCGTCA 60
 ACCGCAGAGG ATGAAGCTGC TGCTGGGCAT CGCCTTGCTG GCCTACGTCG CCTCTGTTTG 120
 55 GGGCAACTTC GTTAATATGA GGTCTATCCA GGAAATGGT GAACTAAAAA TTGAAAGCAA 180
 GATTGAAGAG ATGGTTGAAC CACTAAGAGA GAAATCAGA GATTTAGAAA AAAGCTTTAC 240
 60 CCAGAAATAC CCACCAGTAA AGTTTATATC AGAAAAGGAT CGGAAAAGAA TTTTGAWTAA 300

	CAGGAGGCGC AGKGTTCGTG GGCTCCCATC TKAAC TGACA AACTCATGAT GGACGGCCAC	360
	GAGGTGACCG TGGTGGACAA TTTCTTCACG GGCAGGAAGA GAAACGTGGA GCACTGGATC	420
5	GGACATGAGA ACTTCGAGTT GATTAACCAC GACGTGTGGG AGCCCCCTCTA CATCGAGGTT	480
	GACCAGATAT ACCATCTGGC ATCTCCAGCC TCCCCTCCAA ACTACATGTA TAATCCTATC	540
10	AAGACATTAA AGACCAATAC GATTGGGACA TTAAACATGT TGGGGCTGGC AAAACGAGTC	600
	GGTGCCCGTC TGCTCCTGGC CTCCACATCG GAGGTGTATG GAGATCCTGA AGTCCACCCT	660
	CAAAGTGAGG ATTACTGGGG CCACGTGAAT CCAATAGGAC CTCGGGCTTG CTACGATGAA	720
15	GGCAAACGTG TTGCAGAGAC CATGTGCTAT GCCTACATGA AGCAGGAAGG CGTGGAAGTG	780
	CGAGTGGCCA GAATCTTCAA CACCTTTGGG CCACGCATGC ACATGAACGA TGGGCGAGTA	840
20	GTCAGCAACT TCATCCTGCA GGCCTCCAG GGGGAGCCAC TCACGGTATA CGGATCCGGG	900
	TCTCAGACAA GGGCGTTCCA GTACGTCAGC GATCTAGTGA ATGGCCTCGT GGCTCTCATG	960
	AACAGCAACG TCAGCAGCCC GGTCAACCTG GGAACCCAG AAGAACACAC AATCCTAGAA	1020
25	TTTGCTCAGT TAATTAAAAA CCTTGTGGT AGCGGAAGTG AAATTCAGTT TCTCTCCGAA	1080
	GCCCAGGATG ACCCAGAGAA AAGAAAACCA GACATCAAAA AAGCAAAGCT GATGCTGGGG	1140
30	TGGGAGCCCG TGGTCCCGCT GGAGGAAGGT TTAAACAAAG CAATTCAC TA CTCCGTAAA	1200
	GAACTCGAGT ACCAGGCAA TAATCAGTAC ATCCCCAAAC CAAAGCCTGC CAGAATAAAG	1260
	AAAGGACGGA CTCGCCACAG CTGAACTCCT CACTTTTAGG ACACAAGACT ACCATTGTAC	1320
35	ACTTGATGGG ATGTATTTTTT GGCTTTTTTT TGTGTGCTT TAAAGAAAGA CTTTAACAGG	1380
	TGTCATGAAG AACAACTGG AATTTTCATC TGAAGCTTGC TTTAATGAAA TGGATGTGCC	1440
40	TAAAGCTCC CCTCAAAAAA CTGCAGATT TGCCTGTCAC TTTTGAATC TCTCTTTTA	1500
	TGTAAATAG CGTAGATGCA TCTCTGCGTA TTTTCAAGTT TTTTATCTT GCTGTGAGAG	1560
	CATATGTGT GACTGTGCTT GACAGTTTTA TTTACTGGTT TCTTTGTGAA GCTGAAAAGG	1620
45	AACATTAAGC GGGACAAAAA ATGCCGATTT TATTTATAAA AGTGGGTACT TAATAAATGA	1680
	GTCGTTATAC TATGCATAAA GAAAAAYCCT AGCAGTATTG TCAGGTGGTG GTGCGCCGGC	1740
50	ATTGATTTTA GGCAGATAA AAGAATCTG TGTGAGAGCT TTATGTTTCT CTTTAAATTC	1800
	AGAGTTTTTC CAAGGTCTAC TTTTGAGTTG CAAACTTGAC TTTGAAATAT TCCTGTTGGT	1860
	CATGATCAAG GATATTTGAA ATCACTACTG TGTTTTGCTG CGTATCTGGG GCGGGGCGAG	1920
55	GTGGGGGGC ACAAAGTTAA CATATTCCTG GTTAACCATG GTTAAATATG CTATTTTAAT	1980
	AAAATATTGA	1990
60		

(2) INFORMATION FOR SEQ ID NO: 276:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2436 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
10 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 276:

AACTCGCTT AGCTCTCCAG GGTNAAACGG GTGAGNCCTT AAAAACAGAA GAGAACAAGA 60
15 TTAAAGTCC GTTGCAATGA AAATAACAAA CAATATCAAT GTTTTAATCA AGGATCTCTT 120
CCACATTCTT CCTTCTTATA AGAGCACAGT AACACTATCC TGGAAACCTG TACAAAAGGT 180
TGAGATTGGG CAAAAGAGAG CCAGTGAAGA TACAACCTCA GGTTCACCAC CCAAGAAATC 240
20 TTCAGCAGGA CAAAAAGAG ATGCCAGGCA GATTTATAAC CCTCCAGTG GGAAATATAG 300
CAGCAATTG GGCAACTTTA ATTATGAGCA GAGAGGAGCC TTCAGGGGAA GTAGAGGTGG 360
25 CCGAGGTTGG GGCACACGAG GAAATCGTAG TCGGGGAAGA CTCTACTGAA TAAGACATCA 420
GCATCTCTCA GCATGTGTCAT GAGCTTAATA TACTTAAATT CTACTACTCA TTGGATTGCC 480
GGGGATGTCC CTTTAAACAG ACTGCTGCCT TCAGCTAAAA ACTTAATGTT CTTTATACCT 540
30 TTGTATGTAT GACCTACTTT TGTAACAGAC CATGGTTGTG TCCAAGGTAA AACCACAGTG 600
ATATTTTGG ATGCTTTGTC TGCAATCTTG ACTTGTTTT GCAGTATCAT TATTCAGACT 660
35 TCAAATTGTG AATCTTTTAA ACATCTTGAT AATTGTGTGT TGAGAGCTGT TCATTCTAAA 720
ATGTAATGAA ATTCAGTCTA GTTCTGCTGA TAAAGATCAT CAGTTTGTGA AGGTTACTGA 780
TTTTCCCTCT CCCTCTTAGT TTTTACCCA ATATATGGAG AAGAGTAATG GTCAATCTTA 840
40 ACATTTTGT TTAATTGTTT AATAAAGCTG CTGGGCAGTG GTGCAGCATT CCTACCTAGT 900
GTCATAAAG CAAAATACTT ACATAGCTTT CTAAAAATAT AGGAATGACA TTACATTTTT 960
45 AGGAGAAAGT AAGTTGCTTT GCACCGCCTA CTTAATTCTT TTCCATATAT TGTGATACAA 1020
ACTTTTGAAT ATGGAATCTT ACTAATTGAA TAGAAATGTG TATGTATAAT ATACATACAT 1080
ACATAAGCAT ATATGTGTGT GTGTGTGTGT ATATATATAT ATATGCATGC TGTGAACTT 1140
50 GACTACACAA CATAAATCAC TTTTAAATT CCAGGAACGG GTAGTCTGAC ACGGTGATTA 1200
TCCTTTTGAG GCTGAATCCG TTATTAACCT GTTATTTAGG TTTTACTCC CAGTAGCAAG 1260
55 GGATTCTAAG TTAGTTGCAC TTACATGATT ATTGTATTT AAAACTAAGA ATAAAGGCTG 1320
CATTTTCAAA GATAAATTGG AATTGCTGTT GGTGAAATAA CAACCAAAT ACTGAATCTG 1380
ATGTACATAC AGGTTTCTAC AGGAAGAGAT GGTATAATTT ACAATTTGGA GATTTAATAA 1440
60

CCAGGGCTAC CCAGAAAAAG TGACTTGATA ACATGGTACC AATAAGTAAG GGATGCTCTC 1500
 TCGGTTTGCT TTTGCCACTT TCAAGATTTT AACTTCTCAG GTTATTAAATC AAAATTATTG 1560
 5 TATAAGTTAG CCAATAGAAT TTTTAGGTTA AAACAACAGA TGGGGGGTTT GTGGAGTGTT 1620
 TAATGTCATG GGCATTTTTA GTAGCATAGA CCCTTTGTTC TGCATTTGAA TGTTCGTAT 1680
 10 ATTTTGTGTT CACAGTTAAT CTTCCCTCCC CAAGTTTGCT ATTCAAATCA ACTGCCGTGAA 1740
 TGACATTTCT AGTAGTCTGA TGTATTTTTC TGAGGAATAG TTTGTGATTC CAATGCAGGT 1800
 GTCTTCATTA CCATTACCTC TACACTGCAG AAGAAGCAAA ACTCCTTTAT TAGAATTACT 1860
 15 GCACATGTGT ATGGGGAAAA TAGTTCTGAA AGGCTAGAAT GATACAAGTG AGCAAAAGTT 1920
 GGTGAGCTTG GCTATGGAGT GGTGGCAATA ATCTCTAAAC ATTCCAAAAG ACCATGAGCT 1980
 GAACCTAAAC TCCCTTGGAA TCTGAACAAA GGAATATAAA ATTGCCATTT GAAACTGAC 2040
 20 CAGCTAATCT GGACCTCAGA GATAGATCAG CCAGTGGCCC AAAGCCATTT CAAGTACAGA 2100
 AATTATAGAG ACTACAGCTA AATAAATTG AACATTAAAT ATAATTTTAC CACTTTTGT 2160
 25 CTTTATAAGC ATATTTGTAA ACTCAGAACT GAGCAGAAGT GACTTTACTT TCTCAAGTTT 2220
 GATACTGAGT TGACTGTTC CTTATCCCTC ACCCTTCCCC TTCCCTTTCC TAAGGCAATA 2280
 GTGCACAACT TAGGTTATTT TTGCTTCCGA ATTTGAATGA AAAACTTAAT GCCATGGATT 2340
 30 TTTTCTTTT GCAAGACACC TGTTTATCAT CTTGTTTAAA TGTAATATGC CCCTTATGCT 2400
 TTTGAAATAA ATTTCTTTT GTAATTTTAA AAAAAA 2436

35

(2) INFORMATION FOR SEQ ID NO: 277:

40

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 782 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 277:

GCCACTGACT TCTCCACCC TTCTGTCTCC CCCATAATAG TTTATTTGGT TGGTCTGGAC 60
 50 TCACTTGTTG CCTTTRATTA AATTCCTAAG GGGCCTGAAG AAGACATTTT TACTGCAGAG 120
 GGTAGAGGC ACTTGAGCAA GGGCCCCACA TCCCAACTCT GGGAGTTGTG GTGGGAGGAG 180
 GCATTCTGG GGGATAGGAC CAGACAAGAT AACAGGAGCT CACATGGNAA GCAGAAGCTG 240
 55 TGACAAGTTT AGTAGTCCCA AAATGGGTTA TATCCCTTCC CCCTTTACAT CAGAATCTTG 300
 TGAAATGGGA AAACAACAGA AGGAGGGGAT CAAAGATAGC TGATCTCACA TGCTTCCAG 360
 60 GCAGGGCARG GGTGGGAGTC AAACCGGGT GACAGGTGGG TGGAGAGCCC TGTTTGAGGT 420

5 TGTGGCTGAT CCCTCTCTGG TATTAGTTT TCCCTGGGA GCAGGAAGCC CTAGGAAGAG 480
GGGACTGCAG GGTCCCRGG GGATCTTTCC TCCCTCCCCT GCATGAGGCA GAGGCAAGCT 540
GCCTGCCAAC CCCCTCCCTC AAGGAATGGC CTGCCCAGG AATGCCACC ACACATACCC 600
TCTCTTTTT TTCTAGTCAA ACTCTGTPTT ATTCCTTGGC TTGCCTCCCT CCTTCTCCC 660
10 CTCTCAACCT TTACTTCTGA TTTCTATTTC ATGGAATTG GGATTGAAGT TAAACTACAA 720
CAGTGCCGCC AACACCAAGT CTGTCAGGAA AAAAATACAA AGAAATTTAA CAAAAAAAAA 780
15 AA 782

20 (2) INFORMATION FOR SEQ ID NO: 278:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 961 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 278:

30 GAGTTCGGC TGGAGACCCG TGCTCTGGGC CGGCGCCTTC ACCATGGCCT CGGCAGAGCT 60
GGACTACACC ATCGAGATCC CGGATCAGCC CTGCTGGAGC CAGAAGAACA GCCCCAGCCC 120
AGGTGGGAAG GAGGCAGAAA CTCGGCAGCC TGTGGTGATT CTYTTGGGCT GGGGTGGCTG 180
35 CAAGGACAAG AACCTTGCCA AGTACAGTGC CATCTACCAC AAAAGGGGCT GCATCGTAAT 240
CCGATACACA GCCCCGTGGC ACATGGTCTT CTCTCCGAG TCACTGGGTA TCCCTTCACT 300
TCGTGTTTTG GCCCAGAAGC TGCTCGAGCT GCTCTTTGAT TATGAGATTG AGAAGGAGCC 360
40 CCTGCTCTTC CATGTCTTCA GCAACGGTGG CGTCATGCTG TACCGCTACG TGCTGGAGCT 420
CCTGCAGACC CGTCGCTTCT GCCGCTGCG TGTGGTGGGC ACCATCTTTG ACAGCGCTCC 480
45 TGGTGACAGC AACCTGGTAG GGGCTCTGCG GGCCCTGGCA GCCATCCTGG AGCGCCGGGC 540
CGCCATGCTG CGCCTGTTGC TGCTGGTGGC CTTTGCCCTG GTGGTCGTCC TGTTCACGT 600
CCTGCTTGCT CCCATCACAG CCCTCTTCCA CACCCACTTC TATGACAGGC TACAGGACGC 660
50 GGGCTCTGCG TGGCCGAGC TCTACCTCTA YTCGAGGGCT GACGAAGTAG TCCTGGCCAG 720
AGACATAGAA CGCATGGTGG AGGCACGCCT GGCACGCCGG GTCCTGGCGC GTTCTGTGGA 780
55 TTTCGTGTA TCTGCACAG TCAGCCACCT CCGTGACTION CCTACTTACT ACACAAGCCT 840
CTGTGTCGAC TTCATGCGCA ACTGCGTCCG CTGCTGAGGC CATGCTCCA TCTCAMCTCT 900
GCTCCAGAAA TAAATGCCTG ACAMCTCCCC AAAAAAAAAA AAAAAAAAAA ACTCGAGGGG 960
60

G

961

5

(2) INFORMATION FOR SEQ ID NO: 279:

(i) SEQUENCE CHARACTERISTICS:

10

(A) LENGTH: 1228 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 279:

20

CGCGCTTTGC AGTTCGGTCT CCTGGTGTAC GGCCAACGCC AAGTAGGGGA TTGCGTTCCC	60
TCCAGTCGCA GCCCTATCAG ATTTGGATAT GTCCTTCATA TTTGATTGGA TTTACAGTGG	120
TTTCAGCAGT GTGCTACAGT TTTTAGGATT ATATAAGAAA ACTGGTAAAC TGGTATTTCT	180
TGGATTGGAT AATGCAGGAA AAACAACATT GCTACACATG CTAAAAGATG ACAGACTTGG	240
ACAACATGTC CCAACATTAC ATCCCACTTC CGAAGAACTG ACCATTGCTG GCATGACGTT	300
TACAACTTTT GATCTGGGTG GACATGTTCA AGCTCGAAGA GTGTGGAAAA ACTACCTTCC	360
TGCTATCAAT GGCATTGTAT TTCTGGTGGG TTGTGCAGAC CACGAAAGGC TGTTAGAGTC	420
AAAAGAAGAA CTTGATTAC TAATGACAGA TGAAACCAAT GCTAATGTGC CTATACTGAT	480
TCTTGGGAAT AAGATCGACA GACCTGAAGC CATCAGTGAA GAGAGGTTGC GAGAGATGTT	540
TGGTTTATAT GGTCAGACAA CAGGAAAGGG GAGTATATCT CTGAAAGAAC TGAATGCCCG	600
ACCCITAGAA GTTTTCATGT GTAGTGTGCT CAAAAGACAA GGTTACGGAG AAGGCTTCCG	660
CTGGATGGCA CAGTACATTG ATTAACACAA ACTCACATTG GTTCCAGGTC TCAACGTTCA	720
GGCTTACTCA GAGATTTGAT TGCTCAACAT GCATAACTTG AATTCAATAG ACTTTTGCTG	780
GTTATAAAAC AGATGTTTTT TAGATTATTA ATATTAAATC AACTTAATTT GAATGAGAAT	840
TGAAACTGA TTCAAGTAAG TTTGAGTATC ACAATGTTAG CTTTCTAATT CCATAAAAGT	900
ACTTGGTTTT TACAGTTTAT AATCTGACAT CACCCAGCG CCATTTGTAA AGAGCAACTT	960
TCCAGCAGTA CATTTGAAGC ACTTTTAAAC AACATGAAAC TATAAACCAT ATTTAAAGC	1020
TCATCATGTT AAATTTTTTA TGTACTTTTC TGGAAGTAGT TTTTAAATTT TAGATTATAT	1080
GTCCACCTAT CKTAAGTGTA CAGTTAATAA TTAGCTTATT CAATGATTGC ATGATGCCTT	1140
ACAGTTTTC ATAACTTTTT TTCTTATGCA AACGTCATGC AATAAAACAA ACTCTAATGT	1200
TTGGCAAAAA AAAAAAAAAA AAANTCGA	1228

60

(2) INFORMATION FOR SEQ ID NO: 280:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 1327 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 280:

10 TCTCGGGTCT CGGACAGGT GAGCACCTG ATGAAGGCCA CGGTCCTGAT GCGGCACCTG 60
GGCGGGTGCA GGAGATCGTG GCGCCCTCC GCAAGGGCGS CGGAGACCGG TTACAGGTGA 120
15 TTTCTGATTT TRACATGACC TTGAGCAGGT TTGCATATAA TGGAAAGCGA TGCCCTTCTT 180
CTTACAATAT TCTGGATAAT AGCAAGATCA TCAGTGAGGA GTGTCGGAAA GAGCTCACAG 240
CGCTCCTTCA CCACTATTAC CCAATTGAGA TCGACCCACA CCGGACCGTC AAGGAGAAGC 300
20 TACCTCATAT GGTGGAATGG TGGACCAAAG CGCACAATCT CCTATGTCAG CAGAAGATTC 360
AGAAGTTTCA GATAGCCCAG GTGGTTAGAG AGTCCAATGC AATGCTCAGG GAGGGATATA 420
25 AGACCTTCTT CAACACACTC TACCATAACA ACATTCCCCT TTTCATCTTT TCTGCGGGCA 480
TTGGTGATAT CCTGGAAGAA ATTATCCGAC AGATGAAAGT GTTCCACCCC AACATCCACA 540
TCGTGTCTAA CTACATGGAT TTTAATGAAG ATGGTMTTCT CCAGGGATTT AAGGGCCAGC 600
30 TGATACACAC ATACAACAAG AACAGCTCTG TGTGTGAGAA CTGTGGTTAC TTCCAGCAAC 660
TTGAGGGCAA AACCAATGTC ATCTGCTGG GAGACTCTAT CGGGGACCTC ACCATGGCCG 720
35 ATGGGGTTCC TGGTGTGAG AACATTCTCA AAATTGGCTT CCTGAATGAC AAGGTGGAGG 780
AGCGGCGGGA NCGCTAACAT GGAATCCTAT GACATCGTGC TGGAGAAGGA CGAGACTCTG 840
GATGTGGTCA ACGGGCTACT GCAGCACATC CTGTGCNAG GGGGTCCAGC TGGAGATGCA 900
40 AGGCCCCCTGA AGGCGCAGGC TCCNAAGKCC SCTGCAGGCC GTGGTGAGGA GGGGCGCCTC 960
CCCAGAGTCT GCTCCCCCGT GAACACAGAG CAGAGCCAGG GTGCCAGCA GTGGCTGGGT 1020
45 CCTTCCGCGC CCCTCCGTCC TCCTTTCCCT GAGCACCTTC ATCACCAGAG GCTTGAAGGA 1080
ACCCCGCCAT GTGGCAGGC ACAGGCACTG TTCTTGGTGA ACCTTGGACC ACAGCATGTC 1140
AGTGCTCTAG GGATGTCTA CTCCAGGGAT TTCTCTCAAA ATTTTAAAC ATGGGAAGTT 1200
50 CAAACAAATA TAATGTGTGA AACAGATCAA AATTTTAAAT ATGAAAAAA AGCTGCTCTG 1260
ATTGAGGGGA TGTGGGTCCG GGTAGAACCT GGACCTCTTG GCCTGGGGGC ACATGGGATG 1320
55 CTTCTAG 1327

60 (2) INFORMATION FOR SEQ ID NO: 281:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 799 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 281:

10 TCACCTGCC TACAGCGTGG AGCTCAGATG ACTGCGCCCT CCACGGTCAC TGTGAGCAGG 60
TGGTATTCAC AGCCTGCATG ACCCTCACGG CCAGCCCTGG GGTGTTCCCC GTCAGTGTGT 120
GGCTTTGGCT GAAGCCTAAT TCCACAGCTC CTTGTTTTTT GAGAGAGACT GAGAGAACCA 180
15 TAATCCTTGC CTGCTGAACC CAGCCTGGGC CTGGATGCTC TGTGAATACA TTATCTTGCG 240
ATGTTGGGTT ATTCCAGCCA AAGACATTTT AAGTGCCTGT AACTGATTTG TACATATTTA 300
20 TAAAAATCTA TTCAGAAATT GGTCCAATAA TGCACGTGCT TTGCCCTGGG TACAGCCAGA 360
GCCCTTCAAC CCCACCTTGG ACTTGAGGAC CTACCTGATG GGACGTTTCC ACGTGTCTCT 420
AGAGAAGGAT TCCTGGATCT AGCTGGTCAC GACGATGTTT TCACCAAGGT CACAGGAGCA 480
25 TTGCGTTCGCT GATGGGGTTG AAGTTTGGTT TGGTTCCTGT TTCAGCCCAA TATGTAGAGA 540
ACATTTGAAA CAGTCTGCAC CTTTGATACG GTATTGCATT TCCAAAGCCA CCAATCCATT 600
30 TTGTGGATTT TATGTGTCTG TGGCTTAATA ATCATAGTAA CAACAATAAT ACCTTTTTCT 660
CCATTTTGCT TGCAGGAAAC ATACCTTAAG TTTTTTTTGT TTTGTTTTTG TTTTTTTGTT 720
TTTTGTTTTT CTTTATGAAG AAAAAATAAA ATAGTCACAT TTTTAATACY AAAAAATGGA 780
35 CAAAAAAGT CGAGGGGGG 799

40

(2) INFORMATION FOR SEQ ID NO: 282:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2196 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 282:

50

AAAGACTCTA ACATCCATGA GCTGAACAT GAGCAAGAGC CTACTTGTGC CKSCCAGATG 60
GCTGAGCCCT TCCGTACCTT CCGAGATGGA TGGGTCTCCT ACTACAACCA GCCTGTGTTT 120
55 CTGGCTGGCA TGGGTCTTGC TTTCTTTTAT ATGACTGTCC TGGGCTTTGA CTGCATCACC 180
ACAGGGTACG CCTACACTCA GGGACTGAGT GGTTCATCC TCAGTATTTT GATGGGAGCA 240
TCAGCTATAA CTGGAATAAT GGGAACTGTA GCTTTTACTT GGCTACGTCG AAAATGTGGT 300
60

	TTGGTTCGGA CAGGTCGTGAT CTCAGGATTG GCACAGCTTT CCTGTTTGAT CTTGTGTGTG	360
	ATCTCTGTAT TCATGCCTGG AAGCCCCCTG GACTTGTCCG TTTCTCCTTT TGAAGATATC	420
5	CGATCAAGGT TCATTCAAGG AGAGTCAATT ACACCTACCA AGATACCTGA AATTACAAC	480
	GAAATATACA TGCTAATGG GTCTAATTCT GCTAATATTG TCCCGGAGAC AAGTCCTGAA	540
10	TCTGTGCCCA TAATCTCTGT CAGTCTGCTG TTTGCAGGCG TCATTGCTGC TAGAATCGGT	600
	CTTTGGTCCT TTGATTTAAC TGTGACACAG TTGCTGCAAG AAAATGTAAT TGAATCTGAA	660
	AGAGGCATTA TAAATGGTGT ACAGAACTCC ATGAATATC TTCTTGATCT TCTGCATTT	720
15	ATCATGGTCA TCCTGGCTCC AAATCTGAA GCTTTTGGCT TGCTCGTATT GATTTCAGTC	780
	TCCTTTGTGG CAATGGGCCA CATTATGTAT TTCCGATTTG CCCAAAATAC TCTGGGAAAC	840
	AAGCTCTTTG CTGCGGTCC TGATGCAAAA GAAGTTAGGA AGGAAATCA AGCAAATACA	900
20	TCTGTGTGTT GAGACAGTTT AACTGTGCT ATCCTGTTAC TAGATTATAT AGAGCACATG	960
	TGCTTATTTT GTAATGCAGA ATTCCAATAA ATGGCTGGGT GTTTTGCTCT GTTTTACCA	1020
25	CAGCTGTGCC TTGAGAACTA AAAGCTGTTT AGGAAACCTA AGTCAGCAGA AATPACTGA	1080
	TTAATTTCCC TTATGTTGAG GCATGGAAAA AAAATGGA AAGAAAACT CAGTTTAAAT	1140
	ACGGAGACTA TAATGATAAC ACTGAATTCC CCTATTTCTC ATGAGTAGAT ACAATCTTAC	1200
30	GTAAAAGAGT GGTAGTCAC GTGAATTCAG TTATCATTTG ACAGATCTT ATCTGTACTA	1260
	GAATTCAGAT ATGTCAGTTT TCTGCAAAAC TCACTCTGT TCAAGACTAG CTAATTTATT	1320
35	TTTTTGATC TTAGTTATTT TTAATAACAA ATTCTTCAAG TATGAAGACT AAATTTGAT	1380
	AACTAATATT ATCCTTATTG ATCCTATTGA TCTTAAGTA TTTACATGTA TGTGGAAAA	1440
	CAAAACACTT AACTAGAAAT CTCTAATAAG GTTTATGGTT TAGCTTAAAG AGCACCTTTG	1500
40	TATTTTATT ATCAGATGGG GCAACATATT GTATGAAGCA TATGTAGCAC TTCACAGCAT	1560
	GGTTATCATG TAAGCTGCAG GTAGAAGCAA AGCTGTAAAG TAGATTATC ACACAATGAC	1620
45	TGCATACAGA CTTCAAATAT GTCAATAGTT TGGTCATAGA ACCTAGAAGC CAAAAGCCAC	1680
	ACAGAAGGGC AAGAATCCCA ATTTAACTCA TGTATCATC ATTAGTGATC TGTGTGTAG	1740
	AACATGAGGG TGTAAGCCTT CAGCCTGGCA AGTTACATGT AGAAAGCCCA CACTTGTGAA	1800
50	GGTTTGTGTT TACAAATCAC TTGATTAAAC AACTCAGGT AGAATATTTT TATTTTACT	1860
	GTTTTATACC CAGAAGTTAT TTCTACATTG TTCTACAGCA AGAATATTCA TAAAAGTATC	1920
55	CCTTTCAAAT GCCTTTGAGA AGAATAGAAG AAAAAAGTT TGTATATATT TAAAAAATT	1980
	GTTTTAAAAG TCAGTTTGCA ACATGCTGT ACCAAGATGG TACTTTGCCT TAACCGTTTA	2040
60	TATGCACTTT CATGGAGACT GCAATACGTT GCTATGAGCA CTTTCTTTAT CCTTGGAGTT	2100

TAATCCTTTG CTTTCATCTT CTACAGTATG ACATAATGAT TTGCTATGTT GTAAAACTTT 2160
TGTA AAAAAT TTCTATATAA AATATTGAA ACTTAA 2196

5

(2) INFORMATION FOR SEQ ID NO: 283:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1185 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 283:

GCAGTTAAGG CTTCTGATAA GGAAAGAGAG TCTGAACAGA GCACACACAT CTGGAGCTCC 60
20 AGGAGTGGGG GATGCAGCAT CAGATTCCAT CTTGAATTTT TGCTAAAATA CTTTGTACTC 120
ATAATGGATC TCAACAAAGA TCTGTATTTT ATCTGTGGCT CCATCTTCCC TCTGGGTCAA 180
GTAGATGTTA AGCTGGACCT TGGCAGCCTT CTTAACATGA AGAGATCTAG CTAGACAGAC 240
25 AGACTCCCCC ATTTATGGGA ACAAGAATTC AATTTATTCT CTATTTATAA AACATTTTTT 300
TAAAGTGCCT TGGGTATAAA AATCTAAATG TCTGCGGTGT GATCAGTCAG GAGCACGTAA 360
30 CTATCACTCT TCGCATCCTT TGGTCACTGG GAGATCCTTT GGGGGCTGGG AGGTCCTTCT 420
GTCCCAGGCT AAAGGAAAAG CTTCAAGAAG GTAAGAGCCA CAGAACCCTC GGCAAGAAAG 480
GCCGTCAGG GAGAATGAAT GGTACAGAGA GGAAAGGAAG GAAAGGGGGT GGAACAGAGG 540
35 TAGAAGGCAA GGAAGGGATG CCGCACTGGA GACCGATGGG GACACTCTAA TTGTGCAAGA 600
GGGAGGATCT TCCTTCTTGA ATGCTGAACA CAGCTAGTCT GAACCTTCCT TGGAAAGTCC 660
40 AGCTGTTTGC CCATGCATAG GGCCAACTCT CCCTGCAAAG CAGCAAATGT GGCTTCTATC 720
AGGAAGGAAA AGTATCCATC AGTGTGACAA GAGGTCACCT TCGAACTTGC ATGAACCTCT 780
TGCGCAGCCA CAAAGAGTCC TGGTAGAAGT GAGGATCGCC TAGTCTTACG GCTGTCCGTT 840
45 TATAGAAGTA GCAGTACAAC ACTGCTGCTA GTCTCTGGAA TACAAACAGC ATTTGAAGTC 900
CATCTGTCCA TATGAAGCTG TTGGAGTTT TCCAGCGTAA GTTCATGACC CAGACATGAA 960
50 GGGAGATGCT GAGGGCAAAG TACACAGCTG TCAGGATGAT GGTCCCTTTG AACTTATGGA 1020
ATAGGAGGTT GACCAGGCCA GCTTGAAGA CGAAGGTGTT GAAGAACATG AGGAAAATGA 1080
TGATGATGTT GAAGAGGACT GCAATATCCT GGATGCACTG AGGGAGAGGY TTCTAGTTCC 1140
55 TTTGAATGAG AGCTGTTTCC CTTGCTCTAA GGCAAGCACC TCCAA 1185

60

(2) INFORMATION FOR SEQ ID NO: 284:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 1634 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 284:

10 AGGGAAAGGG GAGGGTAGCG GGAGGGTAGC AGGTGAGTTC CTAGGGCTGG AAGGTTTAGC 60
AGCAGCCTGG TGCAGTGCCC TGTCATCAAG ACAAAACCCAC GGTCCCTMCTG GGTGCCTACC 120
15 AAGCTTG GTT TGTACAAAAG CAAGGTGGGA GTCTATTTT GTACATGAGA TACATCACAC 180
TTACCTGTGG GCCAGTATG TGAAGTGAGT CTGAGTTGTT TACACTGATG CCTTCCCTGC 240
CCACCACAAA TTGTGTACAT AGTCTTCAGA TGATACCACC CCTTTCCCCA GCTCCCAACC 300
20 AAGAGCTGGT TCTAGGCCCTG TGTATATGT CATATTTAGC STTTTATAT ATGACCTTTG 360
ATTTCTGTG TTTGTATTT AGCACAGTG ATGCACCTTC ATTTAAATAC ATCTGTGTGC 420
25 ATACAGATAC GCATATATGT GTGTGCGTAT GCATATATCT CTCATCTGTA GTTTCCAAGA 480
GTTCAGCTGA AGCAGATGGA GTCTGCAGC CCAGGAGACA CCCTGCATCC CTGCTAATAG 540
TGTTTGCCAC AAGTATTAGT GAGTCTTCCT TATTAATATT TTCATTTAG AAGACTGAAG 600
30 CAAAGCTGAT AGTGTGCTG GTTCTTTGG CAGCTAAGTG AGGGTCTTG GATGACTTGC 660
TGTGTTCCCT AAGCTGCACT TTGGGGCCAT CTCTGCAGTA TTAGCCCCCT TTTTGCTTGG 720
35 TGGTACTCTG TCTGTGCCCTG TGTGTGTGTG TGATAGTCAC TCTTGCATGG CTTCATGTC 780
TGGTTTGTGG CATTTGGGGA TAAGGTGCTG AAGCCAGAGC ATTTGCAGTT TGTTTGAGGC 840
CTCGTTGCCA ATGATAGATC ACTCCTGTTG ACCTGGTATG TCTGCTTGCT TGCTGCTTTT 900
40 CCTTGTCTT TCTTGGAAGA GGAAAGGACT CTGGTCAGGC CCAGGCTGAG TGAGATGAGC 960
TGCAGCTGGC TCATGGCCTT CTTAGAGCAG AGAGAGGAGT ATGTCATTTT ACTAAGTTCC 1020
45 TAAACAAACA TTTATGCAGG CAACACTCCT TGCAGATCCA GAAACTGAGG CACAATAGGG 1080
TTATGACTTG CTCAAGAATA TGTAGCTGCT AGGGGGTAAA TCAAGGCATC ACAATTTCTG 1140
TTTCCGGGGC AGGAATAGGC TGTGAATTGC TAGCACTTTT TTTTTTTAAG CAATTACTTT 1200
50 TTGACTTGTT CCTCTGAAAG TGCAAGAGGC GTACACCTTT CCCAAATGTA GACTAGAATC 1260
TGCAGGATGC CACCCACTGT ATAGTTCTGC TTTCCAGAG AGGAAGAACT TTTAGAAACC 1320
55 AAATGATCTT AATTGTTATT GCCACCCCT GGCTTTTCCG GGTAGAAAAT TCACAGTAGG 1380
AATGATTGTT AAGAGAGAGT GCTTGAACC ATGGGTTAAC AGGAAAGGCT ACCTAACTTC 1440
ACATATCTGC AACCAGAGCA GCCACCAAGC ATTACTTAGC AGCAGGAAAA TGATTGTATT 1500
60

TGAGTTCTCG TGTGTCCAAA ACTGAGGCAC CATGTTCTTT GAAACATGC CACCTCAAGG 1560
CTGGGCGCGG TGGCTCACAC CTGTAATCCC AGCAYTTTGG GGAGGCCSAG GGC GGCGGCGG 1620
5 KTTACCGGG GGTG 1634

10 (2) INFORMATION FOR SEQ ID NO: 285:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1795 base pairs
(B) TYPE: nucleic acid
15 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 285:

20 TTCCCCCAG GTTGGCTTCC TTCGATTCCT TTTCTTGGTA TCAACGTTTG ATTGGAAGAA 60
CAACCCCTC TTTGTCAACC TCAATAATGA GCTCACTGTG GAGGAGCAGC TCGGGCACAG 120
CTCMCCGYA TGGTCATTGT TACCCCCCAA GACCGCAAAA ACTCTGTGTG GACACAGGAT 180
25 GGACCTCAG CCCAGATCCT GCAGCAGCTT GTGGTCCTGG CAGCTGAAGC CCTGCCCATG 240
TTAGAGAAGC AGCTCATGGA TCCCCGGGGA CCTGGGGACA TCAGGACAGT GTTCCGGCCG 300
30 CCCTTGGACA TTTACGACGT GCTGATTCGC CTGTYTCTC GCCATATCCC GCGGCACCGC 360
AGGCTTGTGG ACTCGCCAGY TGCTCCTTC TGCGGGGCC TGCTCAGCCA GCCGGGGCCC 420
TCATCCCTGA TGCCCGTGCT GGGTINATGAT CCTNCTCAGC TCTATCTGAC GCAGCTCAGG 480
35 GAGGCTTTG GGGATCTGGC CCTTTCTTC TATGACCAGC ATGGTGGAGA GGTGATTGGT 540
GTCTCTGGA AGCCACACG CTTCAGCCG CAGCCCTTCA AGGCTCCAG CACAAAGGGG 600
40 CGCATGGTGA TGTCTCGAGG TGGGAGCTA GTAATGGTGC CCAATGTTGA AGCAATCCTG 660
GAGGACTTTG CTGTGCTGGG TGAAGGCTG GTGCAGACTG TGGAGGCCG AAGTGAGAGG 720
TGGACTGTGT GATCCAGCT CTGGAGCAAG CTGTAGACGG ACAGCAGGAC APTGGACCTC 780
45 TAGAGCAAGA TGTCAGTAGG ATGACCTCCA CCTCCTTGG ACATGAATCC TCCATGGAGG 840
GCCTGCTGGC TGAACATGCT GAATCATCTC CAACAAAACC CAGCCCCAAC TTTCTCTCTG 900
50 ATGCTCCAGC ATTGGGGCAG GGCATGGTG GCCCATGTAG TCTCTGGGC CTCACCATCC 960
CAGAAGAGGA GTGGGAGCCA GCTCAGAGAA GGAAGTGAAC CCAGGAGATC CATCCACCTA 1020
TTAGCCCTGG GCCTGGACCT CCCTGCGATT TCCCACTCCT TTCTTAGTCT TCTTCCAGAA 1080
55 ACAGAGAAGG GGATGTGTG CTGGGAGAGG CTCTGTCTCC TTCTGTCTG CAGGACCTGT 1140
GCCTAGACTT AGCATGCCCT TCACTGCAGT GTCAGGCCTT TAGATGGGAC CCAGCGAAAA 1200
60 TGTGGCCCTT CTGAGTCACA TCACCGACAC TGAGCAGTGG AAAGGGGCTA TATGTGTATG 1260

AATAGACCAC ATTGAAGGAG CACAATGCCC TCCTGTGTTG ATGCCACTTC CCAGGGTGA 1320
GACACTGGAA AAGAACCGAG GACAGGAAAG GATTGGGTAG GTGAAGGGT CAGGGGACTG 1380
5 GTAGTACCCC AATCTTGGAG AGGTGCAAAA AGCACTGGGG GCTACCCGTT AGCTGCATCT 1440
GCCCCTGGCTG TTTGCCCGTT CATGTCACAA ACTGCCACTA CTATGTACCT GCAGTGGGGT 1500
10 TGCAGAGATG GGGGAGACTC AAGTCTTACT CCCAGGAGC TCCCAGGGCC CAAGGAGGAG 1560
AATGCTGCCT CCTTTCAGTC TGGTCTACAC CCACTTTCCTG GTAGCCTCTC TGCTTCCTGT 1620
AATTCTGGCT GTTTTCCAG ACTCAGCTCA AATAGTGCCC CTCCCTAAGC CCATCCCTCG 1680
15 CCCCAGCCT GAGGTGATCT TTCCCTCCTC TGAATAATTA GAGCAGTTAC TGTCTGTTCA 1740
GTTCTTTGG CAGGCACACA CAGTGGCATA AATTCTATTG TTTTGAATC TGATT 1795
20

(2) INFORMATION FOR SEQ ID NO: 286:

25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 858 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
30
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 286:

TCTGCTTTCG GTGCTGCGTG TACTGCTGGG CGGCTTCTTC GCGCTCGTGG GGTGGCCAA 60
35 GCTCTCGGAG GAGATCTCGG CTCCAGTTTC GGAGCGGATG AATGCCCTGT TCGTGCAGTT 120
TGCTGAGGTG TTCCCGCTGA AGGTATTGG CTACCAGCCA GATCCCCTGA ACTACCAAT 180
AGCTGTGGGC TTTCTGGAAC TGCTGGCTGG GTTGCTGCTG GTCATGGGCC CACCGATGCT 240
40 GCAAGAGATC AGTAACTTGT TCTTGATTCT GTCATGATG GGGGCTATCT TCACCTTGGC 300
AGCTCTGAAA GAGTCACTAA GCACCTGTAT CCCAGCCATT GTCTGCCTGG GGTTCCTGCT 360
45 GCTGCTGAAT GTCGGCCAGC TCTAGCCCA GACTAAGAAG GTGGTCAGAC CCACTAGGAA 420
GAAGACTCTA AGTACATTCA AGGAATCCTG GAAGTAGAGC ATCTCTGTCT CTTTATGCCA 480
TGCAGCTGTC ACAGCAGGAA CATGGTAGAA CACAGAGTCT ATCATCTTGT TACCAGTATA 540
50 ATATCCAGGG TCAGCCAGTG TTGAAAGAGA CATTTTGTCT ACCTGGCACT GCTTCTCTTT 600
TTTAGCTTTA CTACTCTTTT GTGAGGAGTA CATGTTATGC ATATTAACAT TCCTCATGTC 660
55 ATATGAAAT ACAAAATAAG CAGAAAAGAA ATTTAAATCA ACCAAAATTC TGATGCCCCA 720
AATAACCACT TTTAATGCCT TGGTGTAAGT ATACCTCTGA ACTTTTCTCT GTGCCTTTAA 780
ACAGATATAT ATTTTTTTTT AATGAAAATA AAACCATATA TCCTATTTTA TTTCTCTCTT 840
60

TTAAACCTT ATAACTA

858

5

(2) INFORMATION FOR SEQ ID NO: 287:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 915 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 287:

GAATTCGGCA CGAGCGCGGC CATGGCGCTC CTGCTTTCGG TGCTGCGTGT ACTGCTGGGC 60
GGCTCTTCG CGCTCGTGGG GTTGGCCAAG CTCTCGGAGG AGATCTCGGC TCCAGTTTCG 120
20 GAGCGGATGA ATGCCCTGTT CGTGCAGTTT GCTGAGGTGT TCCCGCTGAA GGTATTTGGC 180
TACCAGCCAG ATCCCTGAA CTACCAAATA GCTGTGGGCT TTCTGGAAGT GCTGGCTGGG 240
TTGCTGCTGG TCATGGGCCC ACCGATGCTG CAAGAGATCA GTAAC TTGTT CTGATTCTG 300
25 CTCATGATGG GGGCTATCTT CACCTTGGCA GCTCTGAAAG AGTCACTAAG CACCTGTATC 360
CCAGCCATG TCTGCCTGGG GTTCTGCTG CTGCTGAATG TCGCCAGCT CTAGCCAG 420
30 ACTAAGAAGG TGGTCAGACC CACTAGGAAG AAGACTCTAA GTACATTCAA GGAATCCTGG 480
AAGTAGAGCA TCTCTCTCTC TTTATGCCAT GCAGCTGTCA CAGCAGGAAC ATGGTAGAAC 540
ACAGAGTCTA TCATCTTGTT ACCAGTATAA TATCCAGGT CAGCCAGTGT TGAAAGAGAC 600
35 ATTTTGTCTA CCTGGCACTG CTTTCTCTTT TTAGCTTTAC TACTCTTTG TGAGGAGTAC 660
ATGTTATGCA TATTAACATT CCTCATGTCA TATGAAAATA CAAAATAAGC AGAAAAGAAA 720
40 TTAAATCAA CCAAATTCT GATGCCCAA ATAACCACTT TTAATGCCTT GGTGTAAGTA 780
TACCTCTGAA CTTTTTCTG TGCCCTTAAA CAGATATATA TTTTTTTTWA ATGAAAATAA 840
AACCATATAT CCTATTTTAT TTCCTCCTTT TAAACCTTA TAACTATAA MAAAAAAAAA 900
45 AAAAAAAAAA CTCGA 915

50

(2) INFORMATION FOR SEQ ID NO: 288:

(i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 1517 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 288:

CCTTGTGGCA ACTAGTGGGT CCCCCGGGCT GCAGNAATTC GGGCAGTGGT TCTGNGTCTG 60
AAGATACTCT GAGTTCCTCT GAGAGATCCA AAGGCTCCGG GAGCAGACCC CCAACCCCCA 120
5 AAAGCAGCCC TCAGAAGACC AGGAAGAGCC CTCAGGTGAC CAGGGGTAGC CCTCAGAAGA 180
CCAGCTGTAG CCTCAGAAG ACCAGGCAGA GCCCTCAGAC GCTGAAGCGG AGCCGAGTGA 240
10 CCACCTCACT TGAAGCTTTG CCCACAGGAC AGTGCTGACA GACAAGAGTG GGCACAGTG 300
GAAGCTGAAG TCCTTCCAGA CCAGGGACAA CCAGGGCATT CTCTATGAAG CTGCACCCAC 360
CTCCACCTTC ACCTGTGACT CAGGACCACA GAAGCAAAG TTCTCACTCA AACTGGATGC 420
15 CAAGGATGGG CGCTTGTTC A TGAGCAGAA CTCTTCCAG CGGGCCGCCA AGCCTCTGCA 480
AGTCAACAAG TGGAAGAAGC TGTACTCGAC CCCACTGCTG GCCATCCCTA CCTGCATGGG 540
TTTCGGTGT T CACCAGGACA AATACAGGT CTTCGGTGTTA CCCAGCCTGG GGAGGAGCCT 600
20 TCAGTCGGCC CTGGATGTCA GCCCAAAGCA TGTGCTGTGC AGAGAGGTCT GTGCTGCAGG 660
TGCCCTGCCG GCTGCTGGAT GCCCTGGAGT TCCTCCATGA GAATGAGTAT GTTCATGGAA 720
25 ATGTGACAGC TGAAAATATC TTTGTGGATC CAGAGGACCA GAGTCAGGTG ACTTTGGCAG 780
GCTATGGCTT CGCMTCCGC TATTGCCCAA GTGGCAAACA CGTGGCCTAC GTGGAAGGCA 840
GCAGGAGCCY TCACGAGGGG GACCTTGAGT TTCAITAGCA TGGACCTGCA CAAGGGATGC 900
30 GGGCCCTCCC GCCGCRGYGA CCTCCAGAGC CTGGGYTAMT GCATGCTGAA GTGGYTCTAM 960
GGGTTCCTGC CATGGACAAA TTGCCTTCCA AMAMTGAGGA CATCATGAAG CAAAAACAGA 1020
35 AGTTGCCTTG GGATTCATTT TAATGTAAGC TKGACTTTGT CATGCCAGAA ACAAGGCTCG 1080
GTCACCGTCA GCAGTTTGCA GTTTTCCACC TCCWCCAGT TCCTCCGTGT GGTTGACCCA 1140
GATATCTCCG TTATGCAGCC GCCTCCGGGG GACCACCTCC CTCCTTTGA GTCAGCCACA 1200
40 GACAGCCTAC TTGACGGCCC CGCTGGCCCC CACATTCAC TGAAGTGTG GGATGCCACA 1260
GTGACCCCT CTCAGGCACA GCATGACCTC CTGAAGTCGA GCCTGCTTGC TTTGAACCTA 1320
45 CCAGTTAAAA TCTCTCAAA ATGTTTGGAT ACCGCCATT GGCCCTCAC AGCCACGAGC 1380
TCCCTGACCA GTGTGCGTGT GTGTGTGTGT GTGTGTCTGT GTGTGTGCTT GGGACGGGTG 1440
GGGAGGTCAC CTTTGGGTGT GCGGTGTGCC CCCAGGACCT GTAAGTAATA AAATCTTTAT 1500
50 TTCCAAAAA AAAAAA 1517

55

(2) INFORMATION FOR SEQ ID NO: 289:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3865 base pairs

(B) TYPE: nucleic acid

60

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 289:

5	TGGAGGGGGG GAGCTTCCTT GAGCAGTGGG CCCAGGCCCTG GCCCTCCACA CTTCATTCTC	60
	TGACCTTTCT CTCTCCTCAT TTCGGTGCAT GTCCCTTTCTG CAGCTGCCCTT TCAGCACAGG	120
10	TGGTTCCACT GGGGGCAGCT AACGCTGAGT GACAAGGATG GGAAGCCACA GGTGCATTTT	180
	ACTCAAGTCT TCTCTAGTCA ATGAGGGGCA CCCAGTGCTT CTAGGGCAGG CTGGGTGGTG	240
	GTCCCTTAGG TATCAGCCTC TCTTACTGTA CTCTCCGGGA ATGTTAACCT TTCTATTTTC	300
15	AGCCTGTGCC ACCTGTCTAG GCAAGCTGGC TTCCCCATPG GCCCTGTGG GTCCACAGCA	360
	GCGTGGCTSC CCCCAGGGC CACCGCTTCT TTCTTGATCC TCTTTCTTA ACAGTACTT	420
20	GGGCTTGAGT CTGGCAAGGA ACCTTGCTTT TAGCTTCACC ACCAAGGAGA GAGGTGACA	480
	TGACCTCCCC GCCCCTCAC CAAGGCTGGG AACAGAGGG ATGTGGTGAG AGCCAGGTC	540
	CTCTGGCCCT CTCCAGGGTG TTTTCCACTA GTCCTACTG TCTTCTCCTT GTAGCTAATC	600
25	AATCAATATT CTTCCTTGC CTGTGGGAG TNGGAGAGTG CTGCTGGGTG TACGCTGCAC	660
	CTGCCCCACTG AGTTGGGGAA AGAGGATAAT CAGTGAGCAC TGTTCTGCTC AGAGCTCCTG	720
30	ATCTACCCCA CCCCCTAGGA TCCAGGACTG GGTCAAAGCT GCATGAAACC AGGCCCTGGC	780
	AGCAACCCTG GGAATGGCTG GAGGTGGGAG AGAACCTGAC TTCTCTTTCC CTCTCCCTCC	840
	TCCAACATTA CTGGAATCT ATCCTGTTAG GATCTTCTGA GCTTGTTTCC CTGCTGGGTG	900
35	GGACAGAGGA CAAAGGAGAA GGGAGGGTCT AGAAGAGGCA GCCCTTCTTT GTCCCTCTGGG	960
	GTAAATGAGC TTGACCTAGA GTAAATGGAG AGACCAAAAG CCTCTGATTT TTAATTTCCA	1020
40	TAAAATGTTA GAAGTATATA TATACATATA TATATTTCTT TAAATTTTGT AGTCTTTGAT	1080
	ATGTCTAAAA ATCCATCCC TCTGCCCTGA AGCCTGAGTG AGACACATGA AGAAAAGTGT	1140
	GTTTCATTTA AAGATGTTAA TTAAATGATT GAAACTTGGC TGTGGCTACT GCTTCTTAAT	1200
45	GTGSGGGGGA CAGGGCAGTG GTCTGGGCC ACATTTAGAA GGGAAAATGT TTTGCTGCT	1260
	GCACACATTG GACCCAAGTA TGGGCTCTT CTGCCTAGTA CTGCCAAAGG GACTGTTAAG	1320
50	GTGTCTGTG CATCTTCTAC CCCCCACCC CCATTACGGG TAAAGGRAAC CCCAGACTAG	1380
	GTGAGGGGCC AGCAGCTGCC TCACATTGTG TTCTCTCCTG AGATGGTCCA GCTCACATCC	1440
	AGACACCTTG TTCAGACATT TTATTTGAAT TTATGACAGT GATGGGGATT TGACTGAGAT	1500
55	GCCTTATGGA GAAGTACCCC ACCCTCTATG AAGACAGAAT CACTCTCTGC CATTCAITCT	1560
	GCCTGATGCT AACAACACGC AGCTGATTTA GGGAGTGTCC CAGCCTAGCT GGATCAAGGG	1620
60	AAATTCCAGG AGCCCTGGGG CAGGCCCTGG NCCCCAGTGC CAAGCCTCAG AGTAAGCAGA	1680

	CATTGGGAAA GTTGCCAACC ACTTGGTAGA CCACTAGGTT CTCTGTTTTC CCTTCCCTTT	1740
5	CCTTTTCAAA TCCCACAGTT TCCTGTTGGG GAGAAGCTGT AATTAGCCTA GTCCAGGTAC	1800
	CAGATCCAG CTAGGGGCGC AGCTGNCTTG GATAACTCCA AGAAAACCTG GGCACCACTA	1860
	TTTTTCCAAT TATAAGGACT GTGGCATAAA TTTTAAATG AGTTATATIG AAACCAGATT	1920
10	TCTCCAGCTG CCAAGGAAG AAGGTAGGGC TGGACTCCCT GCTGTGGCCC AGCCCTTGT	1980
	AGGGGTGGT CTCTCACTGC AGCCAGACAG GATGATCCTG GGTCTGGGG AGGGTAAGCT	2040
15	GCCCCTGCC GAGTCTGCA CCGAATAAG AGTCCAAACC CGCTGCTTCC GTGTCTGAG	2100
	AGATGGGTAA ATGGGTGATG GATGGAGCAG ACTGAAGAGA CAGCAGATGA CTCAGTGGTG	2160
	GAAGAAGGGG GGAAGATGCT GGGCTGGCTA GCTAATGTTC CCCCTTTCA GCGATTACA	2220
20	GGAAATGGAG CCCAGCTTGG TCATGAAGTT GGTGTGCTTC CACTGTGCGA TGCCTCTC	2280
	AGAAATTTTG AAGTCAGCCT GCAACTTCTC GAAGACTTTC TTCTTGGGCT TGAGCTCCTC	2340
25	ATCTGGTTGG CCCTTTTCAT AGCCCTTCAC AAACACGTGC TCACCAGGAG CAGAGCCTGC	2400
	CGGAGGGTCC AGAGGTCAA CTGGCGTTT ATCCCTTCTA TAGAAGCACA CAGAAGCATG	2460
	CCTTGGGACT CGACTCCTCT CATCTTCTGG GGTTCAGGT TGCACAGCAC CACTACCAGC	2520
30	CTGTCTGCA GTTCTCTCTT GGGCACGAAC TGTACCAGGC CGCTCACCAC AGTCCGTGGT	2580
	TCAGCTTCCC CCAGTCAAT CTCTCTACA TACAGGCTGT CTGCATCTGG GTGCTTCTCC	2640
35	ACAGTGATGA TTTTCCCAC ACGGATATCC AGCCGGGATG GGATGACCTC CTCTGGTTCT	2700
	GAATTCTTGG CAGGCCTTTG GCCATTGGCT TCTGCTTTGA GGGATCTGGG TAGGCAGCGC	2760
	TGGCCAGTTT TTTCAGGCA GGGGTATTAA ACTTTTCCCG GATTGGATCC AGCAACTTGT	2820
40	TCAGTGCGAC TTCAACAGAA TTCTTCAGGT CTCCAGGATG TACAACCTCA GCAGCAAAGT	2880
	CCTTTTCCAG GTCCACGTAA GCTGTGTAGG TTTTGTTC ACCCCATTTC TCATCTCGTA	2940
45	GGATCAGAAA CTGGACTTA AGGGGAAAAA GGACATGCTT GATGAAGGAC AGAACCCCAT	3000
	TGTCTCCAC ATTCTCTGGC TCACAGAAGG CCTTCTTCAG TTTTCTCTC ACATCCTCCT	3060
	TCCGATCAAG GAGATCAATC TTGGACTCCT CTTCTGAAGA GCTCATTTTG CTGCCTGTTA	3120
50	ATCCTGGAAC CATAGGATTC ATCAGATGGA CCCGTTTGA ATAGCCAAGT GCAGGGAGGT	3180
	ACTTCTCTGC AAAGGTGAAA ATCTTCTCT GATCAATGCC TCCAAATTGG GCATCTACTT	3240
55	TTAAATACTC TTCATCAAA GCCTGCAGTC CGGGGTATAA GAGGCCACTC AGCAAAGGT	3300
	GCTCCACCTG CTTTACCACC TCAGCTCCAG CCTTCTTGA ATCGTGCTGT GTGACCACGG	3360
	AGGAGAGTCT GTACACATCT AGTGTGTACT CTTTGCTGAG CTGGTAATCA GTGCCTTTGA	3420
60	TGAACCTGAG CTTCTCCAAG GGCACACAA TGCTCTCCAG CATTGCTTTG ATCACATTCT	3480

CATAGTAACT GACTCGGAGT TCTAGAAGTT CCCATGGGGC TTTCATGTTA TCCAGGTATG 3540
CGTGGAGGTC CGCAAACAGA ATTGTTACCT CACACCCTGC CTTTAAGAAG TCTGCAATCT 3600
5 TTGACATGGG CACAAAGTAA GCCACATGTG GTTTGCCCCGTT GGTGCGCGTT CCCCAGTAAA 3660
TTTTAAGTTC CCGCTCCTTC AGTATCTCCT TCAGCTTCTC TTCCCCCAGA ACCTCCTGCA 3720
10 GGTTCGCGGT GATAAGGTGC AGTTTCTCTT CAGGGCTGGG AGCGTCCCCC ATGGTCCGCT 3780
ACCCCTGCTT CCCCCTCA GCCCGGCACC AGAGCCCTT CCTGGGTCAC CGTCGCCGCC 3840
GCGTGCCGGG AACTGTCACG CGAGT 3865
15

(2) INFORMATION FOR SEQ ID NO: 290:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1910 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

25

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 290:

AGGGAGAGGA GGAGAGGGG TCTGCGCGG GCCGCTACCC AGAAGCCAGC GGACGGCAGC 60
30 ACGGAGTGGG CTGTCCCGA GCCCAGCCCC GAGCGAGCCC CCCCCCGCC CCGMAGGAC 120
GCGCTYCCA GCCAGCCGA CTCCTAGGAG GAGGGAGGC GGGAAAGCAG CTCAGCCTC 180
35 ACCCACC GCC CTGCCCCAG CCCCCTACT CCCAGGCTCC TCGGACTCG GCGGTCTC 240
CTGGAGTCT CGGAGGGAC CGNCTGTGCA GACGCCATGG AGTTGGTGCT GTCTTCTC 300
TGCAGCTGC TGGCCCCAT GGTCTGGCC AGTGCAGTG AAAAGGAGAA GGAAATGAC 360
40 CCTTTTCATT ATGATTACCA GACCCTGAGG ATTGGGGGAC TGGTGTTCG TGTGTCCTC 420
TTCTCGGTTG GGATCTCTT TATCCTAAGT CGCAGGTGCA AGTGCAGTTT CAATCAGAAG 480
45 CCCCCGCCC CAGGAGATGA GGAAGCCCAG GTGGAGAACC TCATCACCGC CAATGCAACA 540
GAGCCCCAGA AAGCAGAGAA CTGAAGTGCA GCCATCAGGT GGAAGCTCT GGAACCTGAG 600
GCGGCTGCTT GAACCTTTGG ATGCAAATGT CGATGCTTAA GAAAACCGGC CACTTCAGCA 660
50 ACAGCCCTTT CCCCAGGAGA AGCCAAGAAC TTGTGTGTCC CCCACCCTAT CCCCTCTAAC 720
ACCATTCCTC CACCTGATGA TGCAACTAAC ACTTGCCTCC CCACTGCAGC CTGCGGTCT 780
55 GCCCACCTCC CGTGATGTGT GTGTGTGTGT GTGTGTGTGT GACTGTGTGT GTTTGCTAAC 840
TGTGCTCTT GTGGCTACTT GTTGTGGAT GGTATGTGT TTGTTAGTGA ACTGTGGACT 900
CGCTTTCCCA GGCAGGGGCT GAGCCACATG GCCATCTGCT CCTCCCTGCC CCGTGGCCC 960
60

TCCATCACCT TCTGCTCCTA GGAGGCTGCT TGTGCCCCGA GACCAGCCCC CTCCCCTGAT 1020
TTAGGGATGC GTAGGGTAAG AGCACGGGCA GTGGTCTTCA GTCGTCTTGG GACCTGGGAA 1080
5 GGTTTGACAG ACTTTGTCAT CATCTCTCAT GGACTCCTTT CACTCCTTTA AAAAAACCT 1140
TGCTTCCTTA TCCCACCTGA TCCAGTCTG AAGGTCTCTT AGCAACTGGA GATACAAAGC 1200
10 AAGGAGCTGG TGAGCCCAGC GTTGACGTCA GGCAGGCTAT GCCCTTCCGT GGTTAATTTT 1260
TTCCCAGGGG CTTCACGAG GAGTCCCCAT CTGCCCCGCC CCTTCACAGA GCGCCCGGGG 1320
ATTCCAGGCC CAGGGCTTCT ACTCTGCCCC TGGGGAATGT GTCCCCTGCA TATCTTCTCA 1380
15 GCAATAACTC CATGGGCTCT GGGACCCTAC CCCTTCCAAC CTTCCCTGCT TCTGAGACTT 1440
CAATCTACAG CCCAGCTCAT CCAGATGCAG ACTACAGTCC CTGCAATTGG GTCTCTGGCA 1500
GGCAATAGTT GAAGGACTCC TGTTCGGTTG GGGCCAGCAC ACCGGGATGG ATGGAGGGAG 1560
20 AGCAGAGGCC TTTGCTTCTC TGCCTACGTC CCCTTAGATG GGCAGCAGAG GCAACTCCCG 1620
CATCCTTTGC TCTGCCGTGC GGTGGTCAGA GCGGTGAGCG AGGTGGGTTG GAGACTCAGC 1680
25 AGGCTCCGTG CAGCCCTTGG GAACAGTGAG AGGTGAAGG TCATAACGAG AGTGGAAGT 1740
CAACCCAGAT CCCGCCCTC CTGTCTCTG TGTTCCTGCG GAAACCAACC AAACCGTGGC 1800
CTGTGACCCA TTGCTGTCT CTGTATCGTG ATCTATCCTC AACAAACA GAAAAAGGA 1860
30 ATAAATATC CTTTGTTCM TAAAAA AAAAAGG GGGGGGGG 1910

35
(2) INFORMATION FOR SEQ ID NO: 291:

40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3276 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 291:

GCGACCGTCG TTTGAGTCGT CGCTGCCGCT GCCGCTGCCA CTGCCACTGC CACCTCGCGG 60
ATCAGGAGCC AGCGTTGTTT GCCGACGCC TCGCTGCCGG TGGGAGGAAG CGAGAGGGAA 120
50 GCGGCTTGGC GGTTTGTCG CGCTGCTCGC CCACCGCTG GAAGAGCCGA GCCCGGCC 180
AGTCGGTCGC TTGCCACCGC TCGTAGCCGT TACCCGCGGG CCGCCACAGC CCGCGCCCG 240
GAGAGGCGCG CGCCATGGCT TCTGGAGCCG ATTCAAAAGG TGATGACCTA TCAACAGCCA 300
55 TTCTCAAACA GAAGAACCGT CCAATCGGT TAATTGTTGA TGAAGCCATC AATGAGGACA 360
ACAGTGTGGT GTCCTTGTCC CAGCCCAAGA TGGATGAATT GCAGTTGTTT CGAGGTGACA 420
60 CAGTGTGCT GAAAGGAAAG AAGAGACGAG AAGCTGTTTG CATCGTCCTT TCTGATGATA 480

	CTTGTTCTGA TGAGAAGATT CGGATGAATA GAGTTGTTTCG GAATAACCTT CGTGTAACGCC	540
	TAGGGGATGT CATCAGCATC CAGCCATGCC CTGATGTGAA GTACGGCAAA CGTATCCATG	600
5	TGCTGCCCCAT TGATGACACA GTGGAAGGCA TTACTGGTAA TCTCTTCGAG GTATACCTTA	660
	AGCCGTACTT CTTGGAAGCG TATCGACCCA TCCGGAAAGG AGACATTTTT CTTGTCCGTG	720
10	GTGGGATGCG TGCTGTGGAG TTCAAAGTGG TGGAAACAGA TCCTAGCCCT TATTGCATTG	780
	TTGCTCCAGA CACAGTGATC CACTGCGAAG GGGAGCCTAT CAAACGAGAG GATGAGGAAG	840
	AGTCCTTGAA TGAAGTAGGG TATGATGACA TTGTGGCTG CAGGAAGCAG CTAGCTCAGA	900
15	TAAAGGAGAT GGTGGAAGTG CCCCTGAGAC ATCCTGCCCT CTTTAAGGCA ATTGGTGTGA	960
	AGCCTCCTAG AGGAATCCTG CTTTACGGAC CTCTGGAAC AGGAAAGACC CTGATTGCTC	1020
20	GAGCTGTAGC AAATGAGACT GGAGCCTTCT TCTTCTTGAT CAATGGTCCT GAGATCATGA	1080
	GCAAATGGC TGGTGAGTCT GAGAGCAACC TTCGTAAGC CTTTGAGGAG GCTGAGAAGA	1140
	ATGCTCCTGC CATCATCTTC ATTGATGAGC TAGATGCCAT CGCTCCCAA AGAGAGAAAA	1200
25	CTCATGGCGA GGTGGAGCGG CGCATTGTAT CACAGTTGTT GACCCTCATG GATGGCCTAA	1260
	AGCAGAGGGC ACATGTGATT GTTATGGCAG CAACCAACAG ACCCAACAGC ATTGACCCAG	1320
30	CTCTACGGCG ATTTGGTCGC TTTGACAGGG AGGTAGATAT TGAATTCTT GATGCTACAG	1380
	GACGCTTAGA GATTCTTCAG ATCCATACCA AGAACATGAA GCTGGCAGAT GATGTGGACC	1440
	TGGAACAGTA GCCAATGAGA CTCACGGGCA TGTGGGTGCT GACTTAGCAG CCCTGTGCTC	1500
35	AGAGGCTGCT CTGCAAGCCA TCCGCAAGAA GATGGATCTC ATTGACCTAG AGGATGAGAC	1560
	CATTGATGCC GAGGTATGA ACTCTCTAGC AGTTACTATG GATGACTTCC GGTGGGCCCTT	1620
40	GAGCCAGAGT AACCCATCAG CACTGCGGGA AACCCTGGTA GAGGTGCCAC AGGTAACCTG	1680
	GGAAGACATC GGGGGCCTAG AGGATGTCAA ACGTGAGCTA CAGGAGCTGG TCCAGTATCC	1740
	TGTGGAGCAC CCAGACAAAT TCCTGAAGTT TGGCATGACA CCTTCCAAGG GAGTTCTGTT	1800
45	CTATGGACCT CTTGGCTGTG GGAAAACCTT GTTGGCCAAA GCCATTGCTA ATGAATGCCA	1860
	GGCCAACTTC ATCTCCATCA AGGGTCCTGA GCTGCTCACC ATGTGGTTTG GGGAGTCTGA	1920
50	GGCCAATGTC AGAGAAATCT TTGACAAGGC CCGCCAAGCT GCGCCCTGTG TGCTATTCTT	1980
	TGATGAGCTG GATTCTGATT CCAAGGCTCG TGGAGGTAAC ATTGGAGATG GTGGTGGGGC	2040
	TGCTGACCGA GTCATCAACC AGATCCTGAC AGAAATGGAT GGCATGTCCA CAAAAAAAAA	2100
55	TGTGTTTCATC ATTGGCGCTA CCAACCGGCC TGACATCATT GATCCTGCCA TCCTCAGACC	2160
	TGGCCGTCTT GATCAGCTCA TCTACATCCC ACTTCCTGAT GAGAAGTCCC GTGTTGCCAT	2220
60	CCTCAAGGCT AACCTGCGCA AGTCCCCAGT TGCCAAGGAT GTGGACTTGG AGTTCTTGGC	2280

TAAATGACT AATGGCTTCT CTGGAGCTGA CCTGACAGAG ATTTGCCAGC GTGCTTGCAA 2340
GCTGGCCATC CGTGAATCCA TCGAGAGTGA GATTAGGCGA GAACGAGAGA GGCAGACAAA 2400
5 CCCATCAGCC ATGGAGGTAG AAGAGGATGA TCCAGTGCCT GAGATCCGTC GAGATCACTT 2460
TGAAGAAGCC ATGCGCTTTG CGCGCCGTTT TGTCAGTGAC AATGACATTG GGAAGTATGA 2520
10 GATGTTTGCC CAGACCCTTC AGCAGAGTCG GGGCTTTGGC AGCTTCAGAT TCCCTTCAGG 2580
GAACCAGGGT GGAGCTGGCC CCAGTCAGGG CAGTGGAGGC GGCACAGGTG GCAGTGTATA 2640
CACAGAAGAC AATGATGATG ACCTGTATGG CTAAGTGGTG GTGGCCAGCG TGCAGTGAGC 2700
15 TGGCCTGCCT GGACCTTGTT CCCTGGGGGT GGGGGCGCTT GCCCAGGAGA GGGACCAGGG 2760
GTGCGCCAC AGCCTGCTCC ATTCTCCAGT CTGAACAGTT CAGCTACAGT CTGACTCTGG 2820
20 ACAGGGGGTT TCTGTTGCAA AAATACAAAA CAAAAGCGAT AAAATAAAG CGATTTTCAT 2880
TTGGTAGGCG GAGAGTGAAT TACCAACAGG GAATTGGGCC TTGGGCTATG CCATTTCTGT 2940
TGTAAGTTGG GGCAGTGCAG GGGACCTGTG TGGGGTGTGA ACCAAGGCAC TACTGCCACC 3000
25 TGCCACAGTA AAGCATCTGC ACTTGACTCA ATGCTGCCCG AGCCCTCCCT TCCCCCTATC 3060
CAACCTGGGT AGGTGGGTAG GGGCCACAGT TGCTGGATGT TTATATAGAG AGTAGGTTGA 3120
30 TTTATTTTAC ATGCTTTTGA GTTAATGTTG GAAACTAAT CACAAGCAGT TTCTAAACCA 3180
AAAAATGACA TGTGTAAAA GGACAATAAA CGTTGGGTCN AAATGGWRA AAAAAAAAAA 3240
AAAAAAGGGG GGCCCTCTA AAGNNCCANN CTTCGT 3276
35

(2) INFORMATION FOR SEQ ID NO: 292:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1695 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 292:

TTGCAATGGT TGAATTCCCC TCCTCAGGCC AGCCTAGGAG AAGAAGTTCG TAGTCCCAGA 60
50 GGTGAGGCAG GAGGCGGCAG TTTCTGGCGG GTGAGGCGG AGCTGAAGTG ACAGCGGAGG 120
CGGAAGCAAC GGTCCGTGGG GCGGAGAAGG GGGCTGGCCC CAGGAGGAGG AGGAAACCTT 180
55 TCCGAGAAAA CAGCAACAAG CTGAGCTGCT GTGACAGAGG GGAACAAGAT GCGGCGCCG 240
AAGGGAGCCT CTGGGTGAGG ACCCAACTGG GGCTCCCGCC GCTGCTGCTG CTGACCATGG 300
60 CCTTGGCCCG AGGTTGGGG ACCGCTTCGG CTGAAGCATT TGAATCGGTC TTGGGTGATA 360

CGGCGTCTTG CCACCGGGCC TGTCAATTGA CCTACCCCTT GCACACCTAC CCTAAGGAAG 420
AGGAGTTGTA CGCATGTCAG AGAGGTTGCA GGCTGTTTTC AATTGTGTCAG TTTGTGGATG 480
5 ATGGAATTGA CTTAAATCGA ACTAAATTGG AATGTGAATC TGCATGTACA GAAGCATATT 540
CCCAATCTGA TGAGCAATAT GCTTGCCATC TTGGTGGCCA GAATCAGCTG CCATTGCGTG 600
10 AACTGAGACA AGAACAACTT ATGTCCCTGA TGCCAAAAAT GCACCTACTC TTTCTCTTAA 660
CTCTGGTGAG GTCATTCTGG AGTGACATGA TGGACTCCGC ACAGAGCTTC ATAACCTCTT 720
CATGGACTTT TTATCTTCAA GCCGATGACG GAAAAATAGT TATATTCCAG TCTAAGCCAG 780
15 AAATCCAGTA CGCACCACAT TTGAGCAGG AGCCTACAAA TTTGAGAGAA TCATCTCTAA 840
GCAAAATGTC CTATCTGCAA ATGAGAAATT CACAAGCGCA CAGGAATTTT CTTGAAGATG 900
GAGAAAGTGA TGGCTTTTTA AGATGCCTCT CTCTTAACTC TGGGTGGATT TTAACCTACAA 960
20 CTCTTGTCCT CTCGGTGATG GTATGCTTTT GGATTTGTTG TGCAACTGTT GCTACAGCTG 1020
TGGAGCAGTA TGTTCCTCTT GAGAAGCTGA GTATCTATGG TGAATGGAG TTTATGAATG 1080
25 AACAAAAGCT AACAGATAT CCAGCTTCTT CTCTTGTTGG TGTAGATCT AAAACTGAAG 1140
ATCATGAAGA AGCAGGGCCT CTACCTACAA AAGTGAATCT TGCTCATTTT GAAATTTAAG 1200
CATTTTTCTT TTTAAAGACA AGTGTAAATG ACATCTAAAA TTCCACTCCT CATAGAGCTT 1260
30 TTAAATGGT TTCATTGGAT ATAGGCCTTA AGAAATCACT ATAAATGCA AATAAAGTTA 1320
CTCAAACTG TGAAGACTGT ATTTGCTATA ACTTTATGG TATTGTTTTT GTAGTAATTT 1380
35 AAGAGGTGGA TGTTTGGGAT TGTATTATTA TTTTACTAAT ATCTGTAGCT ATTTTGTTTT 1440
TTGCTTTGGT TATTGTTTTT TTCCCTTTTC TTAGCTATGA GCTGATCATT GCTCCTTCTC 1500
ACCTCCTGCC ATGATACTGT CAGTTACCTT AGTTAACAAG CTGAATATTT AGTAGAAATG 1560
40 ATGCTTCTGC TCAGGAATGG CCCACAAATC TGTAATTTGA AATTTAGCAG GAAATGACCT 1620
TTAATGACAC TACATTTTCA GGAAGTAAA TCATTAAAA TTTATTGAA TAATTAAAAA 1680
45 AAAAAAAAAA AANCT 1695

50 (2) INFORMATION FOR SEQ ID NO: 293:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1501 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 293:

60 CACTTTCAGC AGTCCTTTGC TCTCTTTGCT TCTACCTCAA ATAGCCCCAG GAGTGGGCTT 60

	TAGTCTCCAA TATGGAGCAT CTCAGCTTC TCCTGGGGGA TGGGGATTGG GATGGGCAGA	120
5	ATCTGTTTTG GWCTCCGGG TTATTTCCAG TGGGTGTAAA AGCAGAGCTG GGCCTTTCCC	180
	TCTCTTATCC CTGAGGGTGG GTAAGAAGGA CTGTATCTAC ACCTGTTCCT CCCTACCTTC	240
	TCTTTTGTTA GGGAGGCCTC ATTCTAAGTT CCTCAAGAGA GTCCTTGGCT TAAAGCTGTA	300
10	GCAAGGGTGT GCTAGGTGGG GGATTGGAG CAAAACCGTC GAGTAGGCAT GATACTGGTA	360
	TGGAGTGGGC CTGCAAAATC AGACAGAAAT GGCTTGAGAA GCCGCAGGGG AGCATGCCTG	420
15	TCTCTCAGTG ATAGAGTATG GGAGGGACCT CCCTAGCTTG GAAAATGAGA ATTGAAGGGG	480
	TTATGAACAA ATAGGATGCC TAGTTGAGGA TGTTCCTAAA GTTTTGTTCCA ATCTTATCAT	540
	TAGTAGATTT TATAAGCCAC AGAGACAAAC CAGAAACGGA ATAATGTTAC TTTGGATGCT	600
20	TTATTTTTTT GTTCTAGGTG TGGCTTTGTA CATGCAGAAG AATGCTATAT GCTGCACATT	660
	TTGCCTTTAA AGTCITACGA CTTTCCCAT TTTAGTCTAA TGGGAAGATA CAGATGTGCA	720
25	AGTCTGCTTT TTTGTTTTTT GTTATTATTT TTTTTTTTTT GCTCTGTGTT ATGGACATTT	780
	TCAGACATGC ACAGAAGTGG AGAGGATGGT CCTTGGACCC MATGTGTCCA TCACCTAGCT	840
	GCATCACITA TCAGCTATGG TCAACCTGGT TTCATCTGTA TCTCTCTCTT TTCACCTGTA	900
30	TTGTTTATTG AAAATCCAAG ACACTATGCC AATGCAACCG TGACTIONTT GGGAGATTGG	960
	TAGTCTCTTT TGATGGTGAT AGTGATGGGG TGCACATCA TAATCACATC AGGTCTGCTT	1020
35	TTTGCTTTTA ATGTAACTA ATGAAGTTCC AGAGATGGGC CTTAGAAATG TGTTTTAAGA	1080
	ATTAACAAGG AGTCTCAAAA AGAAATGAGA GGGATGCTTC CTTTNCCTTT GCATCTACAA	1140
	AACMAGAGAG AGACTGTTCT GTTGTAAAC TCTTTCAAAA ATTCTGATAT GGTAAGGTAC	1200
40	TTGAGACCTT TCACCAGAAT GTCAATCTTT TTTTCTGTGT AACATGGAAA CTGTGTGAC	1260
	CATTAGCATT GTTATCAGCT TGTACTGGTC TCATAACTCT GGTTTTGGA GAATAATTTG	1320
45	GAAATTGTTG CTGTGTTCTG TGAAAATAAC CTCCCCAAAA TAATTAGTAA CTGGTTGTTT	1380
	TACTTGGTAA TTTGACACCC TGTAAATAAC GCAATTATTT CTGTGTCTTT AAACAGTATA	1440
	AATAGTTGTA AGTTGCAATG CATGATGGAA AAATAAAAAC CTGTATCTCT GTTAAAAAAA	1500
50	A	1501

55 (2) INFORMATION FOR SEQ ID NO: 294:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2683 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

60

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 294:

5	TGANTGTGGT CCCGGGTGCN GATTGGCAGN GCCTCCGCCG CGGCTCGTGG TTGTCCCGCC	60
	ATGGCACTGT CGCGGGGGCT GCCCGGGGAG CTGGCTGAGG CGGTGGCCCG GGGCCGGGTR	120
	CTGGTGGTGG GGGCGGGCGG CATCGGCTGC GAGCTCCTCA AGAATCTCGT GCTCACCGGT	180
10	TTCTCCACA TCGACCTGAT TGATCTGGAT ACTATTGATG TAAGCAACCT CAACAGACAG	240
	TTTTTGTTTC AAAAGAAACA TGTTGGAAGA TCAAAGGCAC AGGTTGCCAA GGAAAGTGTA	300
15	CTGCAGTTTT ACCCGAAAGC TAATATCGTT GCCTACCATG ACAGCATCAT GAACCCTGAC	360
	TATAATGTGG AATTTTCCG ACAGTTTATA CTGGTTATGA ATGCTTTAGA TAACAGAGCT	420
	GCCCGAAACC ATGTTAATAG AATGTGCCGT GCAGCTGATG TTCTCTTAT TGAAAGTGGA	480
20	ACAGCTGGGT ATCTTGACA AGTAACTACT ATCAAAAAGG GTGTGACCGA GTGTTATGAG	540
	TGTCATCCTA AGCCGACCCA GAGAACCPTT CCTGGCTGTA CAATTCGTAA CACACCTTCA	600
25	GAACCTATAC ATTGCATCGT TTGGGCAAAG TACTTGTTCA ACCAGTTGTT TGGGGAAGAA	660
	GATGCTGATC AAGAAGTATC TCCTGACAGA GCTGACCCTG AAGCTGCCTG GGAACCAACG	720
	GAAGCCGAAG CCAGAGCTAG AGCATCTAAT GAAGATGGTG ACATTAAACG TATTTCTACT	780
30	AAGGAATGGG CTAAATCAAC TGGATATGAT CCAGTTNAAA CTTTTTACCA AGCTTTTTAA	840
	AGATGACATC AGGTATCTGT TGACAATGGA CAACTATGG CGGAAAAGGA AACCTCCAAT	900
35	TCCGTTGGAC TGGGCTGAAG TACAAAGTCA AGGAGAAGAA ACGAATGCAT CAGATCAACA	960
	GAATGAACCC CAGTTAGGCC TGAAAGACCA GCAGGTTCTA GATGTAAAGA GCTATGCACG	1020
	TCTTTTTTCA AAGAGCATCG AGACTTTGAG AGTTTATTTA GCAGAAAAGG GGGATGGAGC	1080
40	TGAGCTCATA TGGGATAAGG ATGACCCATC TGCAATGGAT TTTGTACACT CTGCTGCAAA	1140
	CCTCAGGATG CATATTTTCA GTATGAATAT GAAGAGTAGA TTTGATATCA AATCAATGGC	1200
45	AGGGAACATT ATTCTGCTA TTGCTACTAC TAATGCAGTA ATTGCTGGGT TGATAGTATT	1260
	GGAAGGATTG AAGATTTTAT CAGGAAAAAT AGACCAGTGC AGAACAATTT TTTTGAATAA	1320
	ACAACCAAAC CCAAGAAAGA AGCTTCTTGT GCCTTGTCGA CTGGATCCTC CCAACCCCAA	1380
50	TTGTTATGTA TGTGCCAGCA AGCCAGAGGT GACTGTGCGG CTGAATGTCC ATAAAGTGAC	1440
	TGTTCTCACC TTACAAGACA AGATAGTGAA AGAAAAATTT GCTATGGTAG CACCAGATGT	1500
55	CCAAATTGAA GATGGGAAAG GAACAATCCT AATATCTTCC GAAGAGGGAG AGACGGAAGC	1560
	TAATAATCAC AAGAAGTTGT CAGAATTTGG AATTAGAAAT GGCAGCCGGC TTCAAGCAGA	1620
60	TGACTTCCTC CAGGACTATA CTTTATTGAT CAACATCCTT CATAGTGAAG ACCTAGGAAA	1680

	GGACGTTGAA TTTGAAGTTG TTGGTGATGC CCCGGAAAAA GTGGGGSCCA AACAAGCTGA	1740
	AGATGCTGCC AAAAGCATAA CCAATGGGCA GTGATGATGG AGCTCAGCCC TCCACCTCCA	1800
5	CAGCTCAAGA GCAAGATGAC GTTCTCATAG TTGATTCGGA TGAAGAAGAT TCTTCAAATA	1860
	ATGCCGACGT CATGAAGAAG AGAGAAGCCG CAAGAGGAAA TTAGATGAGA AAGAGAATCT	1920
	CAGTGCAAAG AGGTCACGTA TAGAACAGAA GGAAGAGCTT GATGATGTCA TAGCATTAGA	1980
10	TTGAACAGAA ATGCCTCTAA ACAGAACCCT CTTACTATTT AGTTTATCTG GGCAGAACCA	2040
	GATGTGTATG TCCTTTGTTC CAAAGGGAAA AAATTGACAG CAGTGACTTG AAAATGATTG	2100
15	TGCTCCCTTT GAAAGCATTC ATTTTGCTAG AACTGTTAGA CACATGTCAG TATGCTGTAT	2160
	TGAAAGTAGG AATATAGTTT TAAAAACCCT TTGAACAAAG TGTGTGCATA ACCAGTCATG	2220
	AGATAAAACA ACACAATGCA TGTTCCTTTT TTAATGTAAA TACCCTTAGG TATCATTAAT	2280
20	AGTTTCAAAA TATGTGTGTT TAGTAAAGTT GATACCTGGT TATAAATATT ATGCCTTTAT	2340
	TTTTGGCTAG AAGAAGAATT ATTTTGTAGC TAGATCTAAC CATTTTCATA CTCTTAACTG	2400
25	ATTGAAACAG ATTCAAAGAA GTATCGAGTG CTATGCATTG AAACPTGTTT TTAAATGTTA	2460
	GATGGCACTA TGTATATTAA TGTAAAACAA TGTTAATTTA CTCAAGTTTT CAGTTTGTAC	2520
	CGCCTGGTAT GTCTGTGTAA GAAGCCAATT TTTGTGTATT GTTACAGTTT CAGGTTATTT	2580
30	ATATTCGATG TTTTGTAAAA CTCAAATAAC GACTATACTT ATGGACCAAA TAAATGGCAY	2640
	TGCATTCTKG TKAAAAAAN NACAGAAAAA AAAAAAACA AGA	2683
35		

(2) INFORMATION FOR SEQ ID NO: 295:

40	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 1454 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 295:	
	GGACTCGGGG TGGCTCTAAG GGGCAGGGAT AGGGCTGGGG AGCGCCGGCC TGTGGCCCTG	60
50	ACCAGCCCTT TCTCGTGACG GTTCCACCCC GATGCAGGTG GTCACGTGCT TGACCGGGGA	120
	CAGCTACCTG ACGCACTGCT TCCTCCAGCA CCTCATGGTC GTGCTGTCTT CTCTGGAACG	180
	CACGCCCTCG CCGGAGCCTG TTGACAAGGA CTTCTACTCC GAGTTTGGGA ACAAGACCAC	240
55	AGGGAAGATG GAGAACTACG AGCTGATCCA CTCTAGTCGC GTCAAGTTTA CCTACCCACG	300
	TGAGGAGGAG ATTGGGGACC TGACGTTTAC TGTGGCCCAA AAGATGGCTG AGCCAGAGAA	360
60	GGCCCCAGCC CTCAGCATCC TGCTGTACGT GCAGGCCTTC CAGGTGGGCA TGCCACCCCC	420

5 TGGGTGCTGC AGGGGCCCC TGCGCCCCAA GACACTCCTG CTCACCAGCT CCGAGATCTT 480
CCTCCTGGAT GAGGACTGTG TCCACTACCC ACTGCCCGAG TTTGCCAAAG AGCCGCCGCA 540
GAGAGACAGG TACCGGCTGG ACGATGGCCG CCGCGTCCGG GACCTGGACC GAGTGCTCAT 600
GGGTACCAG ACCTACCCGC AGCCCTCACC CTCGTYTTCG ATGACGTGCA AGGTCATGAC 660
10 CTCATGGGCA GTGTACCCCT GGACCACTTT GGGGAGGTGC CAGGTGGCCC GGCTAGAGCC 720
AGCCAGGGCC GTGAAGTCCA GTGGCAGGTG TTTGTCCCA GTGCTGAGAG CAGAGAGAAG 780
CTCATCTCGC TGTGGCTCG CCAGTGGGAG GCCCTGTGTG GCCTGAGCTG CCTGTGAGC 840
15 TCACCGGCTA GCCCAGGCCA CAGCCAGCCT GTCGTGTCCA GCCTGACGCC TACTGGGGCA 900
GGGCAGCAGG CTTTGTGTGT CTCTAAAAAT GTTTTATCCT CCCTTTGGTA CCTTAATTG 960
20 ACTGTCTCG CAGAAATGTG AACATGTGTG TGTGTGTGT TAATTCTTTC TCATGTTGGG 1020
AGTGAGAATG CCGGGCCCCCT CAGGGCTGTT CGGTGTGCTG TCAGCCTCCC ACAGGTGGTA 1080
CAGCCGTGCA CACCAGTGTG GTGTCTGCTG TTGTGGGACC GTTGTTAACA CGTGACACTG 1140
25 TGGGTCTGAC TTTYTCTTCT ACACGTCCCT TCCTGAAGTG TCGAGTCCAG TCCTTTGTTG 1200
CTGTTGCTGT TGCTGTGTCT GTTGCTGTG GCATCTTGCT GCTAATCCTG AGGCTGGTAG 1260
30 CAGAAATGCAC ATTGGAAGCT CCCACCCCAT ATGTGTTCTC AAAGTGGAGG TCTCCCCTGA 1320
TCCAGACAAG TGGGAGAGCC CGTGGGGGCA GGGGACCTGG AGCTGCCAGC ACCAAGCGTG 1380
ATTCTGTCTG CTGTATTCT CTATTCCAAT AAAGCAGAGT TTGACACCGW MAAAAAAAAA 1440
35 AAAAAAAAAA AACN 1454

40 (2) INFORMATION FOR SEQ ID NO: 296:

(i) SEQUENCE CHARACTERISTICS:
45 (A) LENGTH: 828 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 296:

50 ACCCTGGCAT GCCCCACAAA CAGATCACCA GCCAGCTTAC ACAGGCATTA ACTCTCCTCA 60
ATGAGGAAGA ATCATTCACA ACTGAGCAAG ACATTCATAT GATCATTTPAA GGAAGTGTTT 120
55 CCCTTATGTG TTAGCAAGTA TAATCGGCTA ACTCCTAAAT CCCAATGAAT AGTCCTAGGC 180
TGGACAGCAA TGGGCTGCAA TTAGGCAGAT AAAGACATCA GTCCCAGTAA ATGAATCCAT 240
AGACTCATCT AGCACCAACT ACCATTAGCA CTATGTTAGG AGCTGCAAGG CCCCAAAGTA 300
60

GAAGATGTGC ATAATGTCTG CTCTTGTTA GCTCAGGAGA CAATTCAGC ACAGACACTA 360
 CAGTTAACGC TGAAGTCAG CTGCAAGTAA TAGCAWGAAC AGTCAGAAAA ATACCTTATG 420
 5 AGGGGGCAGG GCTGAAGCTG GGCCTTGAAG GATGGATGAA ATTTGGATAG AGAATGAGGA 480
 AGACAGAGGG NCTCCAAGTG AGAGAAGCAT GAAAAATGAG CARGGGCCTG GATCAGTGGG 540
 GTGTATTCAG AGCACCTYTC CAGATGCACC ATGCATGCTC ACAGTCCCTT GCCTATGTGT 600
 10 GGCAGAGTGT CCCAGCCAGA TGTGTGCCCC CACCCCATGT CCATTTACAT GTCCTTCAAT 660
 GCCCACCTCA AAAGGYACYT CTTCTGTAAA GCTTTCCCTK GGTATCAGGA ATCAAAATTA 720
 15 ATCAGGGATC TTTTCACACT GCTGTTTTTT CCTCTTTGGT CCTTCTATCA CTAAAACTCA 780
 TCTCATTCAG CTTACAGCA TAACTAATTA TTTGTTTCC TCACTACA 828

20

(2) INFORMATION FOR SEQ ID NO: 297:

(i) SEQUENCE CHARACTERISTICS:
 25 (A) LENGTH: 2416 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 297:

TCAATTTCCA TTAAGTACA TCAGCCATTG TGATTACCA TTTGTCAGGC TCTCAGGTTT 60
 AACAAAACCT ACTATACCA TCATCCTTCA ACAGCCACAG TCTGAATTGA GCCAACATTT 120
 35 TTTTTTCTTT GAGAAAGAAG TGGACTGGGG CACAACTTTT AGTCTGAGGG GAGCTAGTGG 180
 AAATCTAGAC AATAGAAGTC ATCGATAGCA GCTTTTCCTC AAATGTGTGA CTCCTCAGGG 240
 40 GCTAAACTGC TCTTAGCTTA GAATTATGCT TTAGTAGAGA TCTAGCAGAT AAGTGGGTTA 300
 ATCACTACCA TCCTGTAACT AGTTATATAG CTTCAGACA TGAGGGAGAC ATCAACAGG 360
 GATGGAAGCA ACCCAAGGA TATGCAAGAA GGCATGATG AACCCCTTC CCTCTGGCAG 420
 45 GAGAACAAGG CCAACCAAGG GACAGACTGG AAAGCACTTA GATGTTTAAG GAGGAGAAAG 480
 GGAAGCTTT GACCAGTCTT TGCCTTTTGC CAAGTTCAGC CAGTTCTCCG CTGCTTGCAA 540
 50 CCTCTAGCGC AGTAACATTT GCAGAATTGC AGATTTTCCC CCAGATACTA GGAGGAAAGG 600
 GACTTTGGGG GGTGGGAAG GGTGCTGGT GTTTTAAAAG CATAAGTTAC CTGTTTGCAC 660
 TGTTTTAAGA TAGGAAAAAA AAATAGTGGG CAAGGTGAAC ATCAGACGTA AATTTGTGTG 720
 55 TTTTATTTT GTCATGCTCT TGAAAATGTT TGACCATTG TAGTATACAC AGTGAAACTT 780
 GATTCTCTGT TGCATAAAC ACTATATTTT TTTGAAATG TTAGTGTCCA AAAGCCTCTT 840
 60 CCTCCCTTT CCTTTTCTTA TGTACTTCTT TCATACTTGC TTTACTGATC AGCCAGGCAA 900

5 TAGCCATCCA AGAGCTAGAG CATGAAACAG GGCCCTTTCC AAGTAGGCTC TGGGTGTCCT 960
AAGCCAGCGT GTGCCCTCTG GTTTAGTGAG TGTAATAGAG TCCCTGGCAC CTTTCTTTGC 1020
AAATGAGGCT AACAGACCAG ACTGCAGCAA GTTATCAGAT TCCTCAATCA GATGCACTAG 1080
GAGTGAGGAG CCCAGGGATG GAGGGGGTTC CTGAAGTATT GCAGTTGGCT GTAGTAGCTG 1140
10 AGTTCTTTTC CATGTTACCG AAACGTAGC CAGTTACAGT TTAATCAGGA AAACGGTAGA 1200
TCAATTCAGC CATGGTAGTG CTGGTTGGCA GGGATTGGTA ACGGAGAGAA CTGCTCATCA 1260
GCCAAAACCT AAGCCTTGCC TTTTAGGAGG CCACCAGCAG AGGGACTTGG TCCTCCTTGT 1320
15 CTGGTACTTG TGTACATGCC GGTGACCTGA GGAATCCACT CACACTGGCG AGCAAAAAGG 1380
GAGCAGTGAT TCTCTTTTCT CTCCCCACCC CCTGCCCTTT GTTACCAACA CCAGTTTCCC 1440
20 AGGGGGTACA TGAGTTTCTG AATTTTTTAA AAATGTTTTT GGTTTGGTTT TTCTGGGGAC 1500
TGATAAGTGC TTTAAGCAAT GTCCATACCC CGTCAAGACT CCCAGCTTAG TCATTTTCTT 1560
GTATTTTCTT GTTCACAGTA TTTGTGTGTG TGCTTGTTTT GGCAGCTCAT TTTGGCTGTA 1620
25 TTATATATTG AGTGATGAAT TGATCCTCTT TTTTCCCTAA GGGATATGAA TTGTTTTTCT 1680
TGTTTATAT TCTGCTGTG AATAGCTGGA GCAAACCTGG GGCTGACACG CGTAAGSTAG 1740
30 GGCTGCAAAAR CGAGAAGAGA GCCGGTGGAG TGTACTTGTC CCTGACAGGC TGACCTACCT 1800
GAGTCTCTGA GCTTTTCAGT CCAAATCTTT GCAAGGCTCA AAATGCCACA GAACCTCTCC 1860
TCTTCTCCCC ACTCCCCATG GCAGGGACCG GACCATCCCT ACATGCAACA TGCTGTTTCT 1920
35 CCAGCCCCCTC CCATTGCCAT GGCAAAACAG GTACCTTTGG GGCATGGGGG CATTACATGG 1980
GATGCTGTGT TAATCGACCA CCTAGCCTTC TCTCTCCCCT CCCGTCTTCC CCCAGAATCA 2040
40 CTTCCTAGGA CACCCGAGCT GCTTGCCAG GGTCTGTGTT CCCTGCTAAC TCCAGAGAAG 2100
CATCCCAGGG CTTGTGACA GTCTCTAATT CCCTTCCCTT CTCGTTAAGA ATCATATTGT 2160
ATAGTAGCTT TCAGACCATA CAGTATTCAT TGGGTTACTC CTATTATTAT CAAGTAGCTG 2220
45 GAATTGTGAA GGTGGGAGTA GTTAGATCTT TAGCTTTTAT TCCTTATTTT TTTGTATTAC 2280
TCTCCATGTG TATAAATTAT TGATCATGTT GCTGGCTTTT ATAAACTCTA AGCGAAGGAG 2340
50 GAGCACTGCC TCAGCCTTTG CACATGGTAA TGAAGCACTG TTTTAAATA AAAGRGRGAA 2400
MCMCAAAAA AAAAAA 2416

55

(2) INFORMATION FOR SEQ ID NO: 298:

(i) SEQUENCE CHARACTERISTICS:

60

(A) LENGTH: 545 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 298:

GAATTCGGCA CGAGCCATGC YTGCCCTCTC CTTGATTCTT ACAGTCACCTT TGTGGCTGT 60
TTCTGACTCA GCAGCTACCT GCATTGTGGC CAAAGGATGA CCTATTCCCTT CTCAGGAGGG 120
10 CAAAAATGTG GAATAGTGTG TGTCCATGCC TCTCCTCATG GGCTACCACC TCTGCCACCG 180
TGGTTAATCA GTAACAACCA GGAGAGAAGC TGCTGGAAC TACCTCTGGG AACTCCCTGG 240
15 ATGGTTTGGT GCAGGAATGT AGTAGGCATA CACGTGGTTG CGTGGATCTG GGCCCTCCTG 300
ATGTGAGTAG AGAGGTAAAA GGSCACCATC TCCTTGACCT YTGGGGAAC CATCCACAAA 360
GAAGATGTTT CCAAGATGCT TCTGAAGATT GSCTAAAAAT AGCCGGTTTC CACCCCGTG 420
20 AATGCATCCA TTCTAGAATG CTCCTTCACC AGGACCAGAG AACTGATTTA CAGAAGTGAC 480
ATGAAAACAT TCCATCCCAG AATTGTCANT ACCTCAAATT NAATTCTCTAC CTATTAAAAA 540
25 NAAAA 545

30 (2) INFORMATION FOR SEQ ID NO: 299:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1530 base pairs
35 (B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 299:

40 GGCTCTGCTG GGCATCATAC TTGTCACTGG GTAAACAGTT TGCCCACTTA CCGCAGATGA 60
AGCTGCTTGC CAGGGCTCTC CGGCTCTGTG AGTTTGGGAG GCAGGCATCT TCCAGGAGGC 120
TGGTGGCTGG CCAGGGATGT GTGGGGCCCC GGCAGGGGTG CTGCGCTCCC GTCCAGGTGG 180
45 TTGGGCCCAG GGCTGATCTC CCACCCTGTG GAGCCTGCAT TACTGGAAGG ATCATGCGGC 240
CAGATGATGC CAACGTGGCC GGCAATGTCC ACGGGGGGAC CATCCTGAAG ATGATCGAGG 300
50 AGGCAGGCGC CATCATCAGC ACCCGGCATT GCAACAGCCA GAACGGGGAG CGCTGTGTGG 360
CCGCCCTGGC TCGTGTGAG CGCACCGACT TCCTGTCTCC CATGTGCATC GGTGAGGTGG 420
CGCATGTCAG CGCGGAGATC ACCTACACCT CCAAGCACTC TGTGGAGGTG CAGGTCAACG 480
55 TGATGTCCGA AAACATCCTC ACAGGTGCCA AAAAGCTGAC CAATAAGGCC ACCCTGTGGT 540
ATGTGCCCCCT GTCGCTGAAG AATGTGGACA AGGTCCTCGA GGTGCCTCCT GTTGTGTATT 600
60 CCCGGCANGA GCAGGAGGAG GAGGGCCGGA AGCGGTATGA AGCCAGAAG CTGGAGCGCA 660

5 TGGAGACCAA GTGGAGGAAC GGGGACATCG TCCAGCCAGT CCTCAACCCA GAGCCGAACA 720
CTGTCAGCTA CAGCCAGTCC AGCTTGATCC ACCTGGTGGG GCCTTCAGAC TGCACCTGCG 780
ACGGCTTTGT GCACGGAGGT GTGACCATGA AGCTCATGGA TGAGGTCGCC GGGATCGTGG 840
CTGCACGCCA CTGCAAGACC AACATCGTCA CAGCTTCCGT GGACGCCATT AATTTTCATG 900
10 ACAAGATCAG AAAAGGCTGC GTCATCACCA TCTCGGGACG CATGACCTTC ACGAGCAATA 960
AGTCCATGGA GATCGAGGTG TTGGTGGACG CCGACCCTGT TGTGGACAGC TCTCAGAAGC 1020
GCTACCGGGC CGCCAGTGCC TTCTTCACCT ACGTGTGCT GAGCCAGGAA GGCAGGTCCG 1080
15 TGCTGTGCC CCAGCTGGTG CCCGAGACCG AGGACGAGAA GAAGCGCTTT GAGGAAGGCA 1140
AAGGGCGGTA CCTGCAGATG AAGGCGAAGC GACAGGGCCA CGCGGAGCCT CAGCCCTAGA 1200
20 CTCCTCCTC CTGCCACTGG TGCTTCGAGT AGCCATGGCA ACGGGCCCAG TGTCCAGTCA 1260
CTTAGAAGTT CCCCCCTGG CAAAAACCC AATTACATT GAGAGCTGGT GTGTCTGAA 1320
GTTTTCGTAT CACAGTGTTA ACCTGTACTC TCTCCTGCAA ACCTACACAC CAAAGCTTTA 1380
25 TTTATATCAT TCCAGTATCA ATGCTACACA GTGTGTCCC GAGCGCCGGG AGGCGTTGGG 1440
CAGAAACCTT CGGGAATGCT TCCGAGCAG CTGTAGGGTA TGGGAAGAAC CCAGCACCAC 1500
30 TMATAAGCT GNTGCTTGGC TGGGAAGNA 1530

35 (2) INFORMATION FOR SEQ ID NO: 300:

(i) SEQUENCE CHARACTERISTICS:

40 (A) LENGTH: 997 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 300:

45 AGGTAGTGAG AGACACATTA CACCTAACCA ACAAGAAGAA GGATCCTCCC CTTTATAATT 60
TAACTATGTT TACAGGGAAT GCGTACATTG TGGCTTCCCG AGNATTTTCGT CCAACATGTT 120
TTGAAGAACC CTAAATCCCA ACAACTGATT GAATGGGTAA AAGACACTTA TAGCCCAGAT 180
50 GAACACCTCT GGGCCACCTT TCAGCGTGCA CGGTGGATGC CTGGCTCTGT TCCCAACCAC 240
CCCAAGTACG ACATCTTCAG ACATGACTTC TATTGCCAGG CTGGTCAAGT GGCAGGGTCA 300
55 TGAGGGAGAC ATCGATAAGG GTGCTCCTTA TGCTCCCTGC TCTGGAATCC ACCAGCGGGC 360
TATCTGCGTT TATGGGGCTG GGGACTTGAA TTGGATGCTT CAAAACCATC ACCTGTGGC 420
CAACAAGTTT GACCCAAAGG TAGATGATAA TGCTCTTCAG TGCTTAGAAG AATACCTACG 480
60

TTATAAGGCC ATCTATGGGA CTGAACTTTG AGACACACTA TGAGAGCGTT GCTACCTGTG 540
GGGCAAGAGC ATGTACAAAC ATGCTCAGAA CTGCTGCGGA CAGTGTGGGT GGGAGACCAG 600
5 GGCTTTGCAA TTCGTGGCAT CCTTTAGGAT AAGAGGGCTG MTATTAGATT GTGGGTAAAT 660
AGATCTTTTG CCTTGCAAAT TGCTGCCTGG GTGRATGCTG CTGTTCCTCT CACCCCTAAC 720
CCTAGTAGTT CCTCCACTAA CTTTCTCACT AAGTGAGAAT GAGAACTGCT GTGATAGGGA 780
10 GAGTGAAGGA GGGATATGTG GTAGAGCACT TGATTTCAGT TGAATGCCTG CTGGTAGCTT 840
TTCCATTCTG TGGAGCTGCC GTTCTAATA ATTCCAGGTT TGGTAGCGTG GAGGAGAACT 900
15 TTGATGGAAG GAGAACCTTC CCTTCTGTAC TGTTAACCTA AAAATAAATA GCTCCTGATT 960
CAAAGTAAGG AAAAAAATAA AAAGAAAAAA AACTCGA 997

20

(2) INFORMATION FOR SEQ ID NO: 301:

(i) SEQUENCE CHARACTERISTICS:
25 (A) LENGTH: 2345 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 301:

TTGAGGCCGA CGTAGGGGC CCGGAAGRAA ACTGCGAGGC GAAGGTGACC GGGACCGAG 60
CATTTTCAGAT CTGCTCGGTA GACCTGGTGC ACCACCACCA TGTGGCTGC AAGGCTGGTG 120
35 TGTCTCCGA CACTACCTTC TAGGGTTTTC CACCCAGCTT TCACCAAGGC CTCCTCTGTT 180
GTGAAGAATT CCATCAGGAA GAATCAATGG CTGTTAACAC CTAGCAGGGA ATATGCCACC 240
40 AAAACAAGAA TTGGGATCCG GCGTGGGAGA ACTGGCCAAG AACTCAAAGA GGCAGCATTG 300
GAACCATCGA TGGAAAAAAT ATTTAAAATT GATCAGATGG GAAGATGGTT TGTGCTGGA 360
GGGGCTGCTG TTGGTCTTGG AGCATGTGC TACTATGGCT TGGGACTGTC TAATGAGATT 420
45 GGAGCTATG AAAAGGCTGT AATTGGCCT CAGTATGTCA AGGATAGAAT TCATTCCACC 480
TATATGTACT TAGCAGGAG TATTGGTTTA ACAGCTTTGT CTGCCATAGC AATCAGCAGA 540
50 ACGCTGTTC TCATGAACTT CATGATGAGA GGCTCTTGGG TGACAATTGG TGTGACCTTT 600
GCAGCCATGG TTGGAGCTGG AATGCTGGTA CGATCAATAC CATATGACCA GAGCCCAGGC 660
CCAAAGCATC TTGCTTGGTT GCTACATTCT GGTGTGATGG GTGCAGTGGT GGCTCCTCTG 720
55 ACAATATTAG GGGTCCTCT TCTATCAGA GCTGCATGGT ACACAGCTGG CATGTGGGA 780
GGCCTCTCCA CTGTGCCAT GTGTGCCCC AGTGAAAAGT TTCTGAACAT GGGTGACCCC 840
60 CTGGGAGTGG GCCTGGGTCT CGTCTTTGTG TCCTCATTTG GATCTATGTT TCTTCCACCT 900

ACCACCGTGG CTGGTGCCAC TCTTTACTCA GTGGCAATGT ACGGTGGATT AGTTCTTTTC 960
AGCATGTTCC TTCTGTATGA TACCCAGAAA GTAATCAAGC GTGCAGAAAT ATCACCAGAT 1020
5 TATGGAGTTC AAAAATATGA TCCCATTAAC TCGATGCTGA GTATCTACAT GGATACATTA 1080
AATATATTTA TGCGAGTTGC AACTATGCTG GCAACTGGAG GCAACAGAAA GAAATGAAGT 1140
10 GACTCAGCTT CTGGCTTCTC TGCTACATCA AATATCTTGT TTAATGGGGC AGATATGCAT 1200
TAAATAGTTT GTACAAGCAG CTTTCGTTGA AGTTTAGAAG ATAAGAAACA TGTCATCATA 1260
TTTAAATGTT CCGGTAATGT GATGCCCTCAG GTCTGCCTTT TTTTCTGGAG AATAAATGCA 1320
15 GTAATCCTCT CCCAAATAAG CACACACATT TTCAATTCTC ATGTTTGAGT GATTTTAAAA 1380
TGTMTTGGTG AATGTGAAAA CTAAAGTTTG TGTATGAGA ATGTAAGTCT TTTTCTACT 1440
20 TTAAATTTA GTAGGTTTAC TGAGTAACTA AAATTTAGCA AACCTGTGTT TGCATATTTT 1500
TTTGGAGTGC AGAATATTGT AATTAATGTC ATAAGTGATT TGGAGCTTTG GTAAAGGGAC 1560
CAGAGAGAAG GAGTCACCTG CAGTCTTTTG TTTTTTTAAA TACTTAGAAC TTAGCACTTG 1620
25 TGTATTGAT TAGTGAGGAG CCAGTAAGAA ACATCTGGGT ATTTGGAAAC AAGTGGTCAT 1680
TGTTACATTC ATCTGCTGAA CTTAACAAAA CTGTTCATCC TGAACAGGC ACAGGTGATG 1740
30 CATCTCTCTG CTGTGCTTC TCAGTGTCTC CTTTCCAATA TAGATGTGGT CATGTTTGAC 1800
TTGTACAGAA TGTTAATCAT ACAGAGAATC CTTGATGGAA TTATATATGT GTGTTTACT 1860
TTTGAATGTT ACAAAGGAA ATAACTTTAA AACTATTCTC AAGAGAAAAT ATTCAAAGCA 1920
35 TGAAATATGT TGCTTTTCC AGAATACAAA CAGTATACTC ATGATTGCTA AGTGTTTTTT 1980
TATTTTGCA TATTTATTGA ACTGTCTAAT TGAATACAGC TTGCTCTTGT CACCTCTTCA 2040
40 AGCTTTCAAG CCTTTATAGA AAAGCTTCTT TGTGGCTTAC ACTGGAAATT ATGAAAGCAG 2100
TTTTCTCCT AAGACTTTTG GTTCTCGCA TTGCCTCTCA GACTAAGCAC TAAAAAGCAA 2160
AGCAAAACAG AACTAGTACT GTCTTAATGA AATATATCAA CCCAAAAGTG TAATGAGGAA 2220
45 AATGCTTCAT TAGTTTCCC TAGCAGACTT TTAATCTCTT TAACTGCTA CACCAATTA 2280
TTCTTGAGAC ATTTGTAAGT CCTTGATAC AGAAGAGTTA TATTTAGGAG GNCITTAATG 2340
50 AAGGG 2345

55 (2) INFORMATION FOR SEQ ID NO: 302:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2369 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

60

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 302:

5	TTTTTTTTTT TTTTTTTTTT TTTTINCAAG ATCATGTGTT ATTTATTACT TCAGATAAAA	60
	AGATAGTATA CATATTAGGG AATCCCTTAA AATTCAACTC TAGAGTTATA CACCATCTAG	120
	TACTTTTGCA ATGAATGTTA ACAACAACAA AAAAAATCTC TAAACACCTG AAAGCCCCAC	180
10	TATTAACATG GACTATGGTA ATAAAAAATT TTGACATTTA ATTTGTTCAA CATATAGTAT	240
	TTACATTATG AAACCAATGG TGATGATACA ATAAAGTGAT AAAGAAATAG TAAAAATAAA	300
15	CTTTAAAAAG CAAAGGTTTA TAGTCTGACA ATGCTAATTA TCCTAATTGT ATATAAAAAA	360
	TTAAACATA GAGCTTCTCG TTACAAAATT CTTAATCCTC TGGGTGTGTA TCATTACTTG	420
	CTACCAATTT ACATGCAACA TCTGCTAGGA CTGACATTTG ATTTTMTTCC CCAAGAATGT	480
20	GTGAGTAGAT AAATGACATT TCAGAGCAGA TATTAATTTA CTTGTGGACA GAAAAAGAAA	540
	CTCAAGATTG GTACTGGTCA CAAGCCTCTT CCCAATAGAA ATTATAAAAA CAGTAAGATA	600
25	AAATTTAAAA AAAATCTAAA AAGGGGATGC ATAGGCAAAG AGTACCATAA ATGGCACAGC	660
	TCAAAAAATC CCAGGACCAA TCAGACACAC ATCTTTTCTC TCTCCTTCAG CGACAAGAGG	720
	TCGATTTTGC CATCAAATAA CCATGATTGA AGCAAGCGAG GGGCACCAGG TGTACAACCTG	780
30	ATTAGATCTT GCAAAATACT AAGATGGGAG CAGGGGTGGC CAGAAGAAGG GGTAATTTAT	840
	ATATAATTCA AACTATATAC AGCATAAATG GAATGCAGCC CATCCCAAAC TGGCTCTGTG	900
35	AAACAAATGG ACCTTTATAG TTAAAATTAT AACAAGTGTA ATAATACAAT AGATTTACAT	960
	GGGAAGCAAA ATCCAAGGGA CATTTTATAT TAAGTATTTA CTGTGCTGTT TCAATTTAAA	1020
	AATAATTTTG CTAAGTATAC ATCTCAACTG AAGTCTATGT AAAAAATGTC CTAATAGATA	1080
40	CAGATATTTA CCTTTGGTGA GTTGAAGGCC TTTTGTGAC TTCTGTCTGA ACTGTAGGCA	1140
	GAATGCTAGA TGTACATGCA CATATGGAGA AACTCAAGCT GAGGTCATCC AAAAGCTGTG	1200
45	CGTATGAGGA GGCTGGAGGT ACTTTGAAAG TCAAAGTAGA CCAGAAACCC AAAACAGGTA	1260
	ACAGTGAGGA TGGCAACAGG GAATGGAATG CCAATATGGC AGTAAACTT TTTTAAAAA	1320
	CAGAAAGAGG AAGGCCTCTC GTACCAGCAG AATCCTGTAC ACGTACAAA AAGAAAAAGC	1380
50	CACCCACCAT TTTGTAAAC AGAAGCCAAT TATAGTGTGG GAAAGTACAA ATTACAGAAA	1440
	ACCAGAAGTC AACAGAAGAA AACTACTTGG TTTACTTGAG AGAAAGGAGA ATGGTTCACC	1500
55	CCGAGCAGAG TTAATTGGTG AACGCCGCCA CCACCGCCA CAGAACCTCA TTGGTGTGG	1560
	CCTTCAGACA TTCCACTTCA GGGCTAAGT CGAGAARNTG CCGCACTCTC TTGGTAGCCA	1620
60	AATCATACTG CTCGTCCAGA AGAGGAGCAA AAGCATCTC CAGGACGTCC GAGGCATGAG	1680

CCAGGTAAAT GAGGGCCAGC AAGCGCCTGT CCATGCGGTG AGGGTCATTC ACCCATTTGT 1740
CAAGAACGGC TTCTGTACT TTCTTGATGA GCGCTGCTT AATGTTGTTA TTGGTGAGGG 1800
5 GATGTGTGT CATGTCAAAA AGTAGGAAGT TCTGTTTCTC TGTGTCAAT ACACCCTTTT 1860
CCACCAGGTT TTTAGCTAAT CGTTCCCGTA CATTTCTTAA CTGATAATGC AATTTTAATG 1920
GATTCATGT CTCACCACTA AGTAATTCAA TCCAGTTCTG GACCGTTTCT GGAGGCTGAG 1980
10 TTTCCTAAC ATGCTCAGA GCTTCATCAA GAAGAACATC CCCTGTTGGA GCATCTGACT 2040
TACAGATTAC CTTTCTGTGT AATAGACTTT TACGTCTCAT TCCACAAGCC TCTAGTTGTA 2100
15 ACCTTCTCT CAATGCTAAT TCAATTAACA TACAGCCAGC TAATCCAGAT GATATACAGT 2160
CATTCAAAA TGATGTGTAA ACCTTCGCGG TCCTTGAGGC CCAGCAGGAG CACTTCCTCC 2220
ATCAGGGTCA GCCGCGTTTC CTGGAGTCG CCCTTGTCGT CGTGTCTCTG CTCGTGCGCG 2280
20 CGGCTCTGCG CGTGTCTCT GCTGCTAGCC GCGCCGCGC CCGCCGCCCG CTCCTTGTCG 2340
GCGGCGTTGC GGGAGGCCTC GGTGCGCCG 2369

25

(2) INFORMATION FOR SEQ ID NO: 303:

30

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1181 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 303:

GGGACGTGTG GTTTCAGCTC GTGCGCCTCC CCGTGGGTTT GCGACGTTTA GCGACTATTG 60
40 CGCCTGCGCC ACGCCGCTG CGAGACTGGG GCCGTGGYTG CTGGTCCCGG GTGATGCTAG 120
GCGGCTCCCT GGGCTCCAGG CTGTTGCGGG GTGTAGGTGG GAGTCACGA CGGTTCCGGG 180
CCCGAGGTGT CCGCGAAGGT GCGGCACATG GCGGCAGGG GAGAGCATGG CTCAGCGGAT 240
45 GGTCTGGGTG GACCTGGAGA TGACAGGATT GGACATTGAG AAGGACCAGA TTATTGAGAT 300
GGCCTGTCTG ATAAGTACT CTGATCTCAA CATTTTGCT GAAGTCCCTA ACCTGATTAT 360
50 AAAACAACCA GATGAGTGC TGGACAGCAT GTCAGATTGG TGTAAGGAGC ATCACGGGAA 420
GTCTGGCCTT ACCAAGGCAG TGAAGGAGAG TACAATTACA TTGCAGCAGG CAGAGTATGA 480
ATTCTGTCTC TTGTACGAC AGCAGACTCC TCCAGGGCTC TGTCCACTTG CAGGAAATTC 540
55 AGTTCATGAA GATAAGAAGT TTCTTGACAA ATACATGCCC CAGTTCATGA AACATCTTCA 600
TTATAGAATA ATTGATGTGA GCACTGTAA AGAACTGTGC AGACGCTGGT ATCCAGAAGA 660
60 ATATGAATTT GCACCAAGA AGGCTGCTTC TCATAGGGCA CTGATGACA TTAGTGAAAG 720

5 CATCAAAGAG CTCAGITTTT ACCGAAATAA CATCTTCAAG AAAAAAATAG ATGAAAAGAA 780
GAGGAAAATT ATAGAAAATG GGGAAAATGA GAAGACCGTG AGTTGATGCC AGTTATCATG 840
CTGCCACTAC ATCGTTATCT GGAGGCAACT TCTGGTGGTT TTTTTCCTC ACGCTGATGG 900
CTTGGCAGAG CMCTTCGGTT AACTTGCATC TCCAGATTGA TTAACAAGC AGACAGCACA 960
10 CGAAATACTA TTTTTCCTCT AATATGCTGT TTCCATTATG ACACAGCAGC TCCTTTGTAA 1020
GTACCAGGTC ATGTCCATCC CTGGTACAT ATATGCATTT GCTTTTAAAC CATTCTTTT 1080
GTTTAAATAA ATAAATAAGT AAATAAGCT AGTCTATTG AAATGCAAAA AAAAAAAAAA 1140
15 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA N 1181

20

(2) INFORMATION FOR SEQ ID NO: 304:

(i) SEQUENCE CHARACTERISTICS:

25

- (A) LENGTH: 1537 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 304:

CTTTTGTGT TCCGGCCGAT CCCACCTCTC CTCGACCCTG GACGTCTACC TTCCGGAGGC 60
CCACATCTTG CCCACTCCGC GCGCGGGGCT AGCGCGGGTT TCAGCGACGG GAGCCCTCAA 120
35 GGGACATGGC AACTACAGCG GCGCCGCGG GCGCGCCCG AANATGGAGC TGGCCCGGAA 180
TGGGGAGGGT TCGAAGAAAA CATCCAGGGC GGAGGCTCAG CTGTGATTGA CATGGAGAAC 240
ATGGATGATA CCTCAGGCTC TAGCTTCGAG GATATGGGTG AGCTGCATCA GCGCCTGCGC 300
40 GAGGAAGAAG TAGACGCTGA TGCAGCTGAT GCAGCTGCTG CTGAAGAGGA GGATGGAGAG 360
TTCCTGGGCA TGAAGGGCTT TAAGGGACAG CTGAGCCGGC AGGTGGCAGA TCAGATGTGG 420
45 CAGGCTGGGA AAAGACAAGC CTCAGGGGCC TTCAGCTTGT ACGCCAACAT CGACATCCTC 480
AGACCCTACT TTGATGTGGA GCCTGCTCAG GTGCGAACAG GGCTCCTGGA GTCCATGATC 540
CCTATCAAGA TGGTCAACTT CCCCCAGAAA ATTGCAGGTG AACTCTATGG ACCTCTCATG 600
50 CTGGTCTTCA CTCTGGTTGC TATCCTACTC CATGGGATGA AGACGTCTGA CACTATTATC 660
CGGGAGGGCA CCCTGATGGG CACAGCCATT GGCACCTGCT TCGGCTACTG GCTGGGAGTC 720
55 TCATCCTTCA TTTACTTCCT TGCTACCTG TGCAACGCCC AGATCACCAT GCTGCAGATG 780
TTGGCACTGC TGGGCTATGG CCTCTTTGGG CATTGCAATG TCCTGTTTAT CACCTATAAT 840
ATCCACCTCC ACGCCCTCTT CTACCTCTTC TGGCTGTTGG TGGGTGGACT GTCCACACTG 900
60

CGCATGGTAG CAGTGTGGT GTCTGGACC GTGGCCCCA CACAGCGGT GTCCTCTGT 960
GGCACCTGG CTGCCCTACA CATGCTCTTC CTGCTCTATC TGCATTTTGC CTACCACAAA 1020
5 GTGNTAGAGG GGATCCTGGA CACACTGGAG GGCCCCAACA TCCCGCCCAT CCAGAGGGTC 1080
CCCAGAGACA TCCCTGCCAT GCTCCCTGCT GCTGGGCTTC CCACCACGT CCTCAACGCC 1140
ACAGCCAAAG CTGTTGCGGT GACCTGCAG TCACACTGAC CCCACCTGAA ATTCTTGGCC 1200
10 AGTCCTCTTT CCCGCAGCTG CAGAGAGGAG GAAGACTATT AAAGGACAGT CCTGATGACA 1260
TGTTTCGTAG ATGGGGTTTG CAGCTGCCAC TGAGCTGTAG CTGCGTAAGT ACCTCCTTGN 1320
15 AGCTGTGGCG ACTTCTGAAA GCACAAGGCC AAGAAGCTCT GGCCAGGACT GCAAGGCTCT 1380
GCAGCCAATG CAGAAAATGG GTCAGCTCCT TTGAGAACCC CTCCCCACCT ACCCCTTCCT 1440
TCCTCTTTAT CTCTCCACA TTGTCTTGCT AAATATAGAC TTGGTAATTA AAAAAAAAAA 1500
20 AAAAAAAAAA AAAAAAAAAA AAAAAAGGG GGNCCCC 1537

25

(2) INFORMATION FOR SEQ ID NO: 305:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1493 base pairs
30 (B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 305:

TGCATGCCAA AACCAATGCC TGCCAAACAA AATCTTAGAC ATCCCAATAT AATATGTTAG 60
TTATATTTCT ATTACATCA TTATTGAAAA TACCCAGCTC AGTGCCTGGC TTAATAAATG 120
40 TTTAATTCCT TTACCTACTC TTGCTCTATT TTTTATTTG AAATGGAGAT GAGCAAAATA 180
ACACATTCAT GGCTGAAGCA ATTTTGTGGA CATTTCTTGT TACCAAAAGA TCTATAATCA 240
GGATGATCCT GAGCTGTTCA AACAAGCTGT ATATAAACAG ACAATGAAAC TCTTTGCAGA 300
45 GCTGGAAATT AAAAGGAAAG AGAGAGAAGC CAAAGAGATG CATGAAAGGA AACGACAAAG 360
GGAAGAAGAG ATTGAAGCTC AAGAAAAAGC CAAACGGGAA AGAGAGTGGC AGAAAACTT 420
50 TGAGGAAAGT CGAGATGGTC GTGTGGACAG CTGGCGAAAC TTCCAAGCCA ATACGAAGGG 480
GAAGAAAGAG AAGAAAAATC GGACCTTCCT GAGACCACCG AAAGTAAAAA TGGAGCAACG 540
TGAGTGACCG CCCAAGGTCA CAGGCACAGA ACCTTTCCCC TGCTATCTCC CTTCCTGCTT 600
55 CGAAGGACTC ATTCTTTCCT CCCACTTCCA CCCCAACATA GAGTAGTATT TGCTTTTTAG 660
TCCATTTTGT TTTCAATACG ATTTAATATC GATCAGAGTA ATTCITTTGT ACATTGAAAT 720
60 GAGGGGCTTG GTTTAAAAAA AGACCTTTCC CTCTCCCTGC CCCTAGAACA ACCAGTATTA 780

GAAGGTGCCA CCATTGGTGC TGCCTTCTCT TCCCACAGCC TGTAACCTCAG TGTMTTGTAC 840
TTCACCTGAAT TGTGATGGTT AGAAACTTCG TGGATAGTTT GTGGAAATCA TCCAATTAAA 900
5 CATACTGCTT AAAACAGTGT TGCTGTGACT TCAGAGACAA GCCTGGAAGG GGCACCTTAG 960
GAAGCCCTTT CGCTTCAGTT GCTCGCTTCT GGGTGTGCTC CCTTCGAAGG CCCAGATAAG 1020
10 ACAGGGAACA CTTGTGAGCA CACAGAGCAG CATCTGATGC CCTGTGGTGT TTGGCATGTG 1080
CCCCCTGTCT ACTGACCAAT CAGTGTGGCA TGAGGCCAC GCCACCCAAA CCTTTCACCTT 1140
TCCAAAGAGC TAGCCGTCCT CCACCCAGTA CCATGTCCTA GCCTGTCTGC ATTTGTTAGT 1200
15 GGTAATATTC TTTATGTATA ATAAATTTT ATACCCAAGC CATTGATGTA CTTTTCCTTG 1260
TACTCTCCCT TGTGGGTCCC TTGTCTGGCT TGGCTGAACC CAAAATGCT TTGGGGTTGG 1320
20 ACAGACCTGG CTGAACCTTA GTTCTTCAT CTATGAAATG GGAATATGAA TTACTGCAGC 1380
AGCTTTTAGG GCAGATTGTC CATGGCATAT ACAAGGTAAC TACCATAGTG CTCCTTGGGT 1440
ATTGCCAATA TCCTATTATT TCTGTGTAAT ATGAAGATAC TGATTGTTT GAG 1493
25

(2) INFORMATION FOR SEQ ID NO: 306:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 577 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 306:

AAITCGGCAG AGGNATTATA TACACTATAC TGGCATTAC TGTTCACCC AGCCCGGAAA 60
40 GTCAGAGATG TATATTGGAA AATTACAAC TCCATCTACA TTGGTTCCCA GGACGCTCTC 120
ATAGCACATT ACCCAAGAAT CTACAACGAT GATAAGAACA CCTATATTCG TTATGAACTT 180
45 GACTATATCT TATAATTTTA TTGTATTATT TGTGTTAAT GCACAGCTAC TTCACACCTT 240
AAACTTGCTT TGATTGGTG ATGTAACTT TTAACATTG CAGATCAGTG TAGAACTGGT 300
CATAGAGGAA GAGCTAGAAA TCCAGTAGCA TGATTTTAA ATAACCTGTC TTTGTTTTTG 360
50 ATGTTAAACA GTAAATGCCA GTAGTGACCA AGAACACAGT GATTATATAC ACTATACTGG 420
AGGGATTICA TTTTAAATC ATCTTATGA AGATTAGAA CTCATCCCTT GTGTTTAAAG 480
55 GGAATGTTA ATTGAGAAAT AAACATTTGT GWACAAAATG YTAACAAAAA AAAAAAAA 540
AAAAAAA AAAAAAAA AAAAAAAA AACTCGA 577

(2) INFORMATION FOR SEQ ID NO: 307:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 2860 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 307:

GTGTNGACCG CTCTCNCAAT ATGGCTCCCC CGGGCTGGCA GRWRKTCRGT CWCKRGTGGC 60
TAGCCTGTCC TGACAGGGGA GAGTTAAGCT CCCGTTCCTCC ACCGTGCCCG CTGGCCAGGT 120
15 GGGCTGAGGG TGACCGAGAG ACCAGAACTT GCTTGTCTGA GCTTAGTGCT CAGAGCTGGG 180
GAGGGAGGTT CCGCCGCTCC TCTGCTGTCA GCGCCGGCAG CCCCTCCCGG CTTCACTTCC 240
20 TCCCGCAGCC CCTGCTACTG AGAAGCTCCG GGATCCACAG AGCCGCCACG CCCTGGCCTC 300
AGCCTGCGGG GCTTCCAGTC AGGCCAACAC CGACGCGCAC TGGGGAGGAA GACAGGACCC 360
TTGACATCTC CATCTGCACA GAGGTCCTGG CTGGAACCGA GCAGCCTCCT CCTCCTAGGA 420
25 TGACCTCACC CTCCAGTCTT CCAGTTTTCG GGTGGAGAC ATTAGATGGA GGCCAAGAAG 480
ATGGCTCTGA GCGGACAGA GGAAAGCTGG ATTTTGGGAG CGGGCTGCCT CCCATGGAGT 540
30 CACAGTTCCA GGGCGAGGAC CGGAAATTCG CCCCTTCAGA TAAGAGTCAA CCTCCAATA 600
CCGAAAGGGA ACAGGTGCCA GTCAGCCGGA TCCAAACCGA TTTGACCGAG ATCGGCTCTT 660
CAATGCGGTC TCCCGGGGTG TCCCGGAGGA TCTGGCTGGA CTTCCAGAGT ACCTGAGCAA 720
35 GACCAGCAAG TACCTCACCG ACTTCGGAAA TACACAGAGG GCTCCACAGG TAAGACGGCC 780
TGATGAAGGC TGTGCTGAAA CCTTAAGGAC GGGGTCAATG CCTGCATTCT GCCACTGCTG 840
40 CAGATCGACC GGGACTCTGG CAATCCTCAG CCCCTGGTAA ATGCCCAGTG CACAGATGAC 900
TATTACCGAG GCCACAGCGC TCTGCACATC GCCATTGAGA AAGAGGAGTC TGCAGTGTGT 960
GAAGCTCCTG GTGGAGAATG GGGCCAATGT GCATGCCCGG GTCTGCGGCG ACTTCTTTCCA 1020
45 GAAGGGCCAA GGGACTTGCT TTTATTTCCG TGAGCTACCC CTCTCTTTGG CCGCTTGCAC 1080
CAAGCAGTGG GATGTGGTAA GCTACCTCCT GGAGAACCCA CACCAGCCCG CCAGCCTGCA 1140
50 GGCCACTGAC TCCAGGGCA ACACAGTCCT GCATGCCCTA GTGGATGATC TCGGACAACT 1200
CAGCTGAGAA CATTGCACTG GTGACCAGCA TGTATGATGG GCTCCTCCAA GCTKGGGSCC 1260
SCCYTCTGCC CTACCGTGCA GCTTGAGGAC ATCCGCAACC TGCAGGATCT CACGCCCTCTG 1320
55 AAGCTGGCCG CCAAGGAGGG CAAGATCGAG ATTTTCAGGC ACATCCTGCA GCGGGAGTTT 1380
TCAGGACTGA GCCACCTTTC CCGAAAGTTC ACCGAGTGGT GCTATGGGCC TGTCCGGGTG 1440
60 TCGCTGTATG ACCTGGCTTC TGTGGACAGC TGTGAGGAGA ACTCAGTGCT GGAGATCAAT 1500

GCCTTTCATT GCAAGAGCCC GCACCGACAC CGAATGGTCG TTTTGGAGCC CCTGAACAAA 1560
CTGCTGCAGG CGAAATGGGA TCTGCTCATC CCCAAGTTCT TCTTAAACTT CCTGTGTAAT 1620
5 CTGATCTACA TGTTCATCTT CACCGCTGTT GCCTACCATC AGCCTACCCT GAAGAAGCAG 1680
GCCGCCCTC ACCTGAAAGC GGAGGTGGA AACTCCATGC TGCTGACGG CCACATCCTT 1740
10 ATCCTGCTAG GGGGGATCTA CCTCCTCGTG GGGCCAGCTG TGGTACTTCT GCGGCGCCA 1800
CGTGTTCATC TGGATCTCGT TCATAGACAG CTACTTTGGA AATCCTCTTC CTGTTCCAGG 1860
CCCTGCTTCA CAGTGGTGTG CCAGGTGCTG TGTTCCTGG GCCATCGAGT GGTACCTGCC 1920
15 CCTGCTTG TGCTGCTGG TGGCTGGCT GGCTGAACCT GCTTTACTAA TACACGTGGC 1980
GTTCAGCAC ACAGGCAGTC TACAGTTTCA TGWTCCTGA AGCCCTGGTG AGCCTGAGCC 2040
20 AGGAGGCTTG GCGCCCCGAA GCTCCTACAG GCCCCAATGC CACAGAGTCA GTGCAGCCCA 2100
TGGAGGGACA GGAGGACGAG GGCAACGGG CCCAGTACAG GGTATCCTG GAAGCCTCCT 2160
TGGAGCTCTT CAAATTCACC ATCGGCATGG GCGAGCTGGC CTTCCAGGAG CAGCTGCACT 2220
25 TCCGCGGCAT GTGTGCTGCTG CTGCTGCTGG CCTACGTGCT GCTCACCTAC ATCCTGCTGC 2280
TCAACATGCT CATCGCCCTC ATGAAGCGAA CGTCACAGTG TCGCCACTGA CAGCTGGAGC 2340
30 ATCTGGAAGC TGCAGAAAGC CATCTCTGTC CTGGAGATGG AGAATGGCTA TTGGTGGTGC 2400
AGGAAAAAGC AGCGGGCAGG TGTGATGCTG ACCGTGGCA CTAAGCCCAG ATGGCAGCCC 2460
CGATGAGCGC TGGTGCTTCA GGGTGGAGGA GGTGAAGTGG GCTTCATGGG GAGCAGACGC 2520
35 TGCCTACGCT GTGTGAGGAC CCGTCAGGGG CAGGTGTCCC TCGAACTCTC GAGAACCTG 2580
TCCTGGCTTC CCTTCCAAG GAGGATGAGG ATGGTGCCTC TGAGGAAAAC TATGTGCCCCG 2640
40 TCCAGCTCCT CCAGTCCAAC TGATGGCCCA GATGCAGCAG GAGGCCAGAG GACAGAGCAG 2700
AGGATCTTTC CAACCACATC TGCTGGCTCT GGGGTCCCAG TGAATTCTGG TGGCAAATAT 2760
ATATTTTCAC TAACTCAAAA AAAAAAAAAA AAAAAAAAAA AAAAVGAGGG GGGGCCCGKT 2820
45 ASCCAAWTTC GCCCTATAAG TGAGTGCCWA TTACGATAAA 2860

50

(2) INFORMATION FOR SEQ ID NO: 308:

(i) SEQUENCE CHARACTERISTICS:

55

- (A) LENGTH: 876 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 308:

CTGCTTGTGT CTGCGCTGGT GCTGGGCTGG CTGAACCTGC TTTACTATAC ACGTGGCTTC 60
CAGCACACAG GCATCTACAG TGTCTGATC CAGAAGCCCT GGTGAGCCTG AGCCAGGANN 120
5 TTGGCGCCCC GAAGCTCCTA CAGGCCCCAA TGCCACAGAG TCAGTGCAGC CCATGGAGGG 180
ACAGGAGGAC GAGGGCAACG GGGCCAGTA CAGGGGTATC CTGGAAGCCT CCTTGGAGCT 240
10 CTTCAAATTC ACCATCGGCA TGGGCGAGCT GGCCTTCCAG GAGCAGCTGC ACTTCCGCGG 300
CATGGTGTCTG CTGCTGTGTC TGGCTACGT GCTGCTCACC TACATCCTGC TGCTCAACAT 360
GCTCATCGCC CTCATGNAGC GAGACCGWCA ACAGTGTGCG CACTGACAGC TGGAGCATCT 420
15 GGAAGCTGCA GAAAGCCATC TCTGTCTGG AGATGGAGAA TGGCTATTGG TGGTGCAGGA 480
AGAAGCAGCG GGCAGGTGTG ATGCTGACCG TTGGCACTAA GCCAGATGGC AGCCCCGATG 540
AGCGCTGGTG CTTCAGGGTG GAGGAGGTGA ACTGGGCTTC ATGGGAGCAG ACGCTGCCTA 600
20 CGCTGTGTGA GGACCCGTCA GGGCAGGTG TCCCTCGAAC TCTCGAGAAC CCTGTCTCTG 660
CTTCCCCCTCC CAAGGAGGAT GAGGATGGTG CCTCTGAGGA AAATATGTG CCCGTCCAGC 720
25 TCCTCCAGTC CAACTGATGG CCCAGATGCA GCAGGAGGCC AGAGGACAGA GCAGAGGATC 780
TTTCCAACCA CATCTGTCTG CTCTGGGCTC CCAGTGAATT CTGGTGGCAA ATATATATTT 840
30 TCACTAAMMM AAAAAAAAAA AAAAAAAAAA ACTCGA 876

(2) INFORMATION FOR SEQ ID NO: 309:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2025 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 309:

CATGACCCGC CTGATGCGAT CCCGCACAGC CTCTGGTTCC AGCGTCACTT CTCTGGATGG 60
45 CACCCGCAGC CGCTCCCA CAAGCGAGGG CACCCGAAGC CGCTCCCA CAAGCGAGGG 120
CACCCGCAGC CGCTCCCA CAAGCGAGGG GGCCACCTG GACATCACC CCAACTGGG 180
50 TGCTGCTGGG AACASGCCG GCCCAAGTCC ATGGAGGTCT CCTGCTAGGC GGCTGCCCA 240
GCTGCCGCCC CCGACTCTG ATCTCTGTAG TGGCCCCCTC CTCCCCGCC CCTTTTCGCC 300
CCCTGCCTGC CATACTGCGC CTAAGTCGGT ATTAATCCAA AGCTTATTTT GTAAGAGTGA 360
55 GCTCTGGTGG AGACAAATGA GGTCTATTAC GTGGGTGCCC TCTCCAAAG CGGGGTGGCG 420
GTGGACCAA GGAAGGAAGC AAGCATCTCC GCATCGCATC CTCTTCCATT AACCAGTGGC 480
60 CGGTTGCCAC TCTCCTCCCC TCCCTCAGAG ACACCAAAT GCCAAAAACA AGACGCGTAC 540

AGCACACACT TCACAAAGCC AAGCCTAGGC CGCCCTGAGC ATCCTGGTTC AAACGGGTGC 600
CTGGTCAGAA GGCCAGCCGC CCACTTCCCG TTTCCTCTTT AACTGAGGAG AAGCTGATCC 660
5 AGTTTCCGGA AACAAAATCC TTTTCTCATT TGGGGAGGGG GGTAAATAGTG ACATGCAGGC 720
ACCTCTTTTA AACAGGCAAA ACAGGAAGGG GGAAAAGGTG GGATTCATGT CGAGGCTAGA 780
10 GGCATTTGGA ACAACAAATC TACGTAGTTA ACTTGAAGAA ACCGATTTTT AAAGTTGGTG 840
CATCTAGAAA GCTTTGAATG CAGAAGCAAA CAAGCTTGAT TTTTCTAGCA TCCTCTTAAT 900
GTGCAGCAAA AGCAGGCRAC AAAATCTCCT GGCTTTACAG AAAAAATAT TTCAGCAAAC 960
15 GTTGGGCATC ATGGTTTPTG AAGGCTTPTG TTCTGCTTTC TGCCCTCTCCT CCACAGCCCC 1020
AACCTCCAC CCCTGATACA TGAGCCAGTG ATTATCTTG TTCAGGGAGA AGATCATTTA 1080
20 GATTTGTTTT GCATTCCTTA GAAATGAGGG CAACATTCCA CAGCTGCCCT GGCTGTGATG 1140
AGTGTCTTG CAGGGGCCGG AGTAGGAGCA CTGGGGTGGG GGCGGAATG GGGTTACTCG 1200
ATGTAAGGA TTCTTGTTG TGTGTTGAG ATCCAGTGCA GTTGTGATT CTGTGGATCC 1260
25 CAGCTTGTT CCAGGAATTT TGTGTGATG GCTTAAATCC AGTTTTCAAT CTTGACAGC 1320
TGGGCTGGA CGTGAATCA GTAGCTGAAC CTGTCTGACC CGGTCACGTT CTGGATCCT 1380
30 CAGAACTCTT TGCTCTTGTG GGGGTGGGG TGGGAATCA CGTGGGAGC GGTGGCTGAG 1440
AAATGTAAG GATTCTGGA TACATATTCC ATGGGACTTT CCTTCCCTCT CTGCTTCCT 1500
CTTTCTCTG TCCCTAACCT TTCGCCAAT GGGGCAGCAC CACTGACGTT TCTGGCGGC 1560
35 CAGTGCGGCT GCCAGGTCC TGTACTACTG CCTTGACTT TTCATTTTG CTCACCGTG 1620
ATTTCTCAT AGGAAGTTG GTCAGAGTGA ATTGAATATT GTAAGTCAGC CACTGGGACC 1680
40 CGAGGATTC TGGGACCCG CAGTTGGGAG GAGGAAGTAG TCCAGCCTTC CAGGTGGCGT 1740
GAGAGGCAAT GACTCGTTAC CTGCCGCCA TCACCTTGA GGCTTCCCT GGCTTGAGT 1800
AGAAAAGTC GGGATCGGG CAAGAGAGGC TGAGTACGGA TGGGAACTA TTGTGCACAA 1860
45 GTCTTTCCAG AGGAGTTCT TAATGAGATA TTTGTATTTA TTTCCAGACC AATAAATTG 1920
TAACCTTGCA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAACTC 1980
50 GAGGGGGCC CGTACCCAAT TCGCCGTATA TGATCGTAAA CAATC 2025

55 (2) INFORMATION FOR SEQ ID NO: 310:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3026 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

60

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 310:

5	TAGGCAGCAC TGAAATATCC TAACCCCTA AGCTCCAGGT GCCCTGTGGN ACGAGCAACT	60
	GGACTATAGC AGGGCTGGGC TCTGTCTTCC TGGTCATAGG CTCACTCTTT CCCCCAAATC	120
	TTCTCTGGA GCTTTGCAGC CAAGGTGCTA AAAGGAATAG GTAGGAGACC TCTTCTATCT	180
10	AATCCTTAAA AGCATAATGT TGAACATTCA TTCAACAGCT GATGCCCTAT AACCCCTGCC	240
	TGGATTCTT CCTATTAGGC TATAAGAAGT AGCAAGATCT TTACATAATT CAGAGTGGTT	300
15	TCATTGCCTT CCTACCTCT CTAATGGCCC CTCCATTTAT TTGACTAAAG CATCACACAG	360
	TGGCACTAGC ATTATACCAA GAGTATGAGA AATACAGTGC TTTATGGCTC TAACATTACT	420
	GCCTTCAGTA TCAAGGCTGC CTGGAGAAAG GATGGCAGCC TCAGGGCTTC CTTATGTCCT	480
20	CCACCACAAG AGCTCCTTGA TGAAGGTCAT CTTTTCCTCC TATCCTGTTC TCCCCCTCCC	540
	CGCTCCTAAT GGTACGTGGG TACCCAGGCT GGTCTCTGGG CTAGGTAGTG GGGACCAAGT	600
25	TCATTACCTC CCTATCAGTT CTAGCATAGT AAACCTACGT ACCAGTGTTA GTGGGAAGAG	660
	CTGGGTTTTC CTAGTATACC CACTGCATCC TACTCCTACC TGGTCAACCC GCTGCTTCCA	720
	GGTATGGGAC CTGCTAAGTG TGAATTACC TGATAAGGGA GAGGGAAATA CAAGGAGGGC	780
30	CTCTGGTGTT CCTGGCCTCA GCCAGCTGCC CACAAGCCAT AAACCAATAA AACAAGAATA	840
	CTGAGTCAGT TTTTATCTG GGTCTCTTTC ATTCCCACTG CACTTGGTGC TGCTTTGGCT	900
35	GACTGGGAAC ACCCCATAAC TACAGAGTCT GACAGGAAGA CTGGAGACTG TCCACTTCTA	960
	GCTCGGAACT TACTGTGTAA ATAACTTTC AGAACTGCTA CCATGAAGTG AAAATGCCAC	1020
	ATTTTGCTTT ATAATTTCTA CCCATGTTGG GAAAAACTGG CTTTTCCTCA GCCCTTTCCA	1080
40	GGGCATAAAA CTCAACCCCT TCGATAGCAA GTCCCATCAG CCTATTATTT TTTTAAAGAA	1140
	AACTTGCACT TGTTTTCTT TTTACAGTTA CTTCTTCTCT GCCCCAAAAT TATAAACTCT	1200
45	AAGTGTA AAAAGTCTTA ACAACAGCTT CTGCTTGTA AAAATATGTA TTATACATCT	1260
	GTATTTTAA ATTCTGCTCC TGAAAAATGA CTGTCCCAT TCCACTCAC TGCATTTGGG	1320
	GCCTTTCCCA TTGGTCTGCA TGTCTTTTAT CATTGCAGGC CAGTGGACAG AGGGAGAAGG	1380
50	GAGAACAGGG GTCGCCAACA CTTGTGTTGC TTTCTGACTG ATCCTGAACA AGAAAGAGTA	1440
	ACACTGAGGC GCTCGCTCCC ATGCACAACT CTCCTCAACA CTTATCCTCC TGCAAGAGTG	1500
55	GGCTTTCCAG GGTCTTTACT GGAAGCAGT TAAGCCCCCT CCTACCCCT TCCTTTTTC	1560
	TTTCTTTACT CCTTTGGCTT CAAAGGATTT TGGAAAAGAA ACAATATGCT TTACTCAT	1620
60	TTTCAATTC TAAATTTGCA GGGGATACTG AAAAATACGG CAGGTGGCCT AAGGCTGCTG	1680

TAAAGTTGAG GGGAGAGGAA ATCTTAAGAT TACAAGATAA AAAACGAATC CCCTAAACAA 1740
AAAGAACAAT AGAACTGGTC TTCCATTTTG CCACCTTTCC TGTTCATGAC AGCTACTAAC 1800
5 CTGGAGACAG TAACATTTC A TTAACCAAAG AAAGTGGGTC ACCTGACCTC TGAAGAGCTG 1860
AGTACTCAGG CCACTCCAAT CACCCTACAA GATGCCAAGG AGGTCCCAGG AAGTCCAGCT 1920
CCTTAAACTG ACGCTAGNMA ATAAACCTGG GCAAGTGAGG CAAGAGAAAT GAGGAAGAAT 1980
10 CCATCTGTGA GGTGAYAGGC AAGGATGAAA GACAAAGAAG GAAAAGAGTA TCAAAGGCAG 2040
AAAGGAGATC ATTTAGTTGG GTCTGAAAGG AAAAGTCTTT GCTATCCGAC ATGTACTGCT 2100
15 AGTACCTGTA AGCATTTTAG GTCCAGAAT GGAAAAAAA ATCAGCTATT GGTAATATAA 2160
TAATGTCCTT TCCCTGGAGT CAGTTTTTTT AAAAAGTTAA CTCTTAGTTT TTA CTGTGTTT 2220
AATTCATAAA GAGAAGGGAG CTGAGGCCAT TCCCTGTAGG AGTAAAGATA AAAGGATAGG 2280
20 AAAAGATTCA AAGCTCTAAT AGAGTCACAG CTTTCCCAGG TATAAAACCT AAAATTAAGA 2340
AGTACAATAA GCAGAGGTGG AAAATGATCT AGTTCCTGAT AGCTACCCAC AGAGCAAGTG 2400
25 ATTTATAAAT TTGAAATCCA AACTACTTTC TTAATATCAC TTTGGTCTCC ATTTTTCCTCA 2460
GGACAGGAAA TATGTCCCCC CCTAACTTTC TTGCTTCAAA AATTAAAAATC CAGCATCCCA 2520
AGATCATCTT ACAAGTAATT TTGCACAGAC ATCTCCTCAC CCCAGTGCCT GTCTGGAGCT 2580
30 CACCCAAGGT CANCCAACA ACTTGGTTGT GAACCAACT GCCTTAACCT TCTGGGGGAG 2640
GGGGATTAGC TAGACTAGGA GACCCAGAAG TGAATGGGAA AGGGTGAGGA CTTCAACAATG 2700
35 TTGGCCTGTC AGAGCTTGAT TAGAAGCCAA GACAGTGGCA GCAAAGGAAG ACTTGGCCCA 2760
GGAAAAACCT GTGGGTGTG CTAATTTCTG TCCAGAAAAT AGGGTGACA GAAGCTTGTG 2820
GGGTGCATGG AGGAATTGGG ACCTGGTTAT GTTGTATATC TCGGACTGTG AATTTTGGTG 2880
40 ATGTAAACA GAATATTCTG TAAACCTAAT GTCTGTATAA ATAATGAGCG TTAACACAGT 2940
AAAATATTCA ATAAGAAGTC AAAAAAAAAA AAAAAAACT CGAGGGGGG CCCGGTACCC 3000
45 AATTTNCCAA ATAGAGATNG TATTAC 3026

50 (2) INFORMATION FOR SEQ ID NO: 311:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 712 base pairs

(B) TYPE: nucleic acid

55 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 311:

60 GCAGGCTTTG TGCTCACCTA CAAGCTGGGT GAGCAGGGTG CCAGCAGCCT GTTTCCTCTT 60

CTCTGCTGG ACCACGGCGT TTCTGCTCCC GAGTTGGGAC TGTGGAATGG TGTGGGTGCT 120
GTGGTCTGCT CCATCGCTGG CTCTCCCTG GGTGGGACCT TGCTGGCCAA GCACTGGAAA 180
5 CTGCTGCCTC TGTTGARGTC GGTGCTGCGC TTCCGCCTCG GGGGCCTAGC CTGTCAGACT 240
GCCTTGGTCT TCCACCTGGA CACCCTGGGG GCCAGCATGG ACGCTGGCAC AATCTTGAGA 300
10 GGGTCAGCCT TGCTGAGCCT ATGTCTGCAG CACTTCTTGG GAGGCCTGGT CACCACAGTC 360
ACCTTCACTG GGATGATGCG CTGCAGCCAG CTGGCCCCCA GGGCCTGCAG GCCACACACT 420
ACAGCCTTCT GCCACGCTG GAGCTGCTGG GGAAGCTGCT GCTGGGCACT CTGCCGAGGC 480
15 CTGGCTGATG GGTGGGGCC ACATCCCTGC TTCTTGCTCC TGCTCATCCT CTCTGCCTTT 540
CCCGTTCTGT ACCTGGACCT AGCAGCCAGC ACCTTCTCTT GAGCTGAGTG GCTGGAGTGG 600
20 TCAATAAAGC CACATGTGCC TGTGGCCCAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 660
AACTGGAGGG GGGGCCCGGT ACCCAAATCG CCGGATATGA TCGTAAACAA TC 712

25

(2) INFORMATION FOR SEQ ID NO: 312:

(i) SEQUENCE CHARACTERISTICS:
30 (A) LENGTH: 1289 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 312:

CAAAATTTCAG GAAGGCAAGA GAATATCAAA CAAAGATTTC TGGAAGTATT 60
TTGCCAACCT TCTGGTTGAG CTGCAAGAAA ATATTTATGG TGAGAACTTT TCTGTTTCCC 120
40 GTTATTGGGT TTTTGGTTGG TTTTGTGTTG TTTTITACTA TGCTTTGGTC TGTAATAATA 180
TGCAACTGAA CTACATTTCAG AAGGAAATAT TGTCTACATA GAATATTATA TGAAGTTGGT 240
45 ACATAATTCCT GATGAGGAAA AAAAATCTTT GCAATTCCTT AAGCCATATT GTTGTPTTTC 300
TGTGTTGTTT TCCCTGGATG AAAATATCAG TATTAAGTAG ACAGCATATT ATTCAAGTGT 360
TTAGACTTAT TAATATGTTT TTGCTCTGTA TTTATACATA TGTGTATTTT GGAAAGTATT 420
50 GCCTTTPTTA AGGGAAGCTA TAATTCGATA CATAGTGAAA AAGGGAATGG TGACCCCTTT 480
GTGCCTCTTC CACTGAGGAT AACAAACAGC ATTGTAATCC ATTCTCTTGC ACCTTCTTCT 540
55 TCTTATCTTG TTATTACGGT TTTATTAATT TTGTAGAGGG ACAGGGAGTG GGCAAGGGGA 600
AGAAGCAGCT TATTTGACTA ACCAGCCCCT CTGTGGTCCA CCAGCGTCTT GGCTTGGTGG 660
GAGGCTCTC AATCAGCAGG GCCCCAGGAG GGAAGAAGAA GTGGGCAAAA GCCTGGCCTC 720
60

GCCGCTCGGG AGCTTTGCCA TCTGAGCCAC GCCTCCTCCA GGCCATGCTC CTTGAACTTG 780
 GAAATGTCAA CCGGAGCCCT TACACCAGCC CTCCAGCATC TAATAGACTT GAATCTACTC 840
 5 TAAACGAATA TTTAATCCAA CCTCACTACA TTGTAGCTCA GTCCAACGAC TAACCCTGAA 900
 ATGGGGGTGT TCCAGCCTTC AGCGAGATGG CCAAGCGGTC CCTGGGGGGC TGTGGCAGCG 960
 GGCTTATCCT TCTCTGTTGC CAACCTTGCC GTCCGACCTC CTCCGCCCCC ATGCGGTGAC 1020
 10 CCCGTCCGTG TCTGTGTCTG TCCATACGTG TGAGTCCAGC TAAAAAGACA AAACAGAACC 1080
 CGTGGGCCCCA GCTCGGAAGG TCGTGGAGA AGGCTCCGAC GTCTCCGAAG TGCAGCCCTT 1140
 15 GGGATGGCAT TCCGTGTGT GCCTTATTCC TGGAGAATCT GTATACGGCT CGCCTATAGA 1200
 AATATAGCCT CTTTATGCTG TATTAAAAGG ACTTTTAAAA GCAAAAAAAA AAAAAAAAAA 1260
 CTTGAGGGGG GGNCCGGTAC CCAATTNTC 1289
 20

25 (2) INFORMATION FOR SEQ ID NO: 313:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 313:

Met Phe Leu Ile Phe Val Tyr Phe Leu Lys Ile Leu Phe Ser Ser Ser
 1 5 10 15

35 Leu Pro Phe Leu Trp Leu
 20

40 (2) INFORMATION FOR SEQ ID NO: 314:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 128 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 314:

Met Met Phe Leu Thr Gln Gly Gly Pro Leu Pro Ser Thr Arg Ala Arg
 1 5 10 15

50 Pro Thr Cys Gln Ala Gly Ala Leu Pro Lys Pro Ser Gly Leu Leu Gly
 20 25 30

55 Val Thr Cys Trp Asn Gly Leu Lys Gly Pro Leu Cys Gly Asn Arg Cys
 35 40 45

Ser Pro Asn Thr Leu Leu Leu Ala Ala Arg Gln Ala Leu Trp Lys Gly
 50 55 60

60 Arg Gly Arg Thr His Gln Asp Leu Pro Gly Pro Leu Gln Gly Arg Gln

[illegible]

15

(2) INFORMATION FOR SEQ ID NO: 315:

20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 315:

25 Met Gln Phe Leu Leu Thr Ala Phe Leu Leu Val Pro Leu Leu Ala Leu
1 5 10 15

30 Cys Asp Val Pro Ile Ser Leu Gly Phe Ser Pro Ser
20 25

(2) INFORMATION FOR SEQ ID NO: 316:

35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 64 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 316:

Met Asp Gly Phe Ser Ser Arg Leu Phe Ser Ser Leu Pro Phe Val Ala
1 5 10 15

Leu Gln Trp Phe Ile Val Ile Ser His Leu Leu Ser Leu Ser Leu Ser
45 20 25 30

Ala Cys Cys Tyr Gln Thr His Cys Ser Leu Xaa Gln Leu Ser Ser Ala
35 40 45

50 Phe Ser Xaa Met Gly Glu Ser Cys Val Gly Glu Arg Glu Tyr Xaa Phe
50 55 60

55

(2) INFORMATION FOR SEQ ID NO: 317:

60 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 317:

5

Met Pro Leu Ile Asn Leu Leu Leu Leu Tyr Tyr Val Pro Asn Gly Gly
 1 5 10 15

10

Lys Gln Asp Lys Lys
 20

15

(2) INFORMATION FOR SEQ ID NO: 318:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 318:

Met Gly Arg His Leu Val Leu Val Met Phe Ile Thr Thr Ser Leu His
 1 5 10 15

25

Ser Gly Thr Pro Val Pro Glu Asn Val Ile Cys Gly Val Thr Lys Gly
 20 25 30

30

Pro Gln Gly Lys Lys Lys Lys
 35

35

(2) INFORMATION FOR SEQ ID NO: 319:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 319:

Met Leu Trp Trp Ser Arg Asp Tyr Thr Met Val Phe Leu Leu Phe Thr
 1 5 10 15

45

Met Val Phe Thr Gly Asp Leu Val Ile Arg Gly Arg Thr Glu Leu Ser
 20 25 30

50

Leu

(2) INFORMATION FOR SEQ ID NO: 320:

55

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 88 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 320:

Met Val Cys Ser Ser Leu Cys Asp Ile Gly Gly Ile Ile Thr Pro Phe

1 5 10 15
 Ile Val Phe Arg Leu Arg Glu Val Trp Gln Ala Leu Pro Leu Ile Leu
 20 25 30
 5 Phe Ala Val Leu Gly Leu Leu Ala Ala Gly Val Thr Leu Leu Leu Pro
 35 40 45
 10 Glu Thr Lys Gly Val Ala Leu Pro Glu Thr Met Lys Asp Ala Glu Asn
 50 55 60
 Leu Gly Arg Lys Ala Lys Pro Lys Glu Asn Thr Ile Tyr Leu Lys Val
 65 70 75 80
 15 Gln Thr Ser Glu Pro Ser Gly Thr
 85

20 (2) INFORMATION FOR SEQ ID NO: 321:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 321:

Met Gln Pro Gly Ala Gly Val Leu Val Leu Gly Leu Leu Leu Pro Pro
 1 5 10 15
 30 Pro Gln Ser Pro Ser Leu Ser
 20

35

(2) INFORMATION FOR SEQ ID NO: 322:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 322:

Met Thr Phe Thr Leu Gly Asp Ser Gln Val Leu Leu Ile Asn Leu Phe
 1 5 10 15
 45 Pro Ser Met Pro Ser Gly Ser Cys Ala Arg Pro
 20 25

50

(2) INFORMATION FOR SEQ ID NO: 323:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 64 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 323:

60 Met Cys Leu Glu Cys Trp Ala Glu Asn Leu Gly Pro His His Thr Ser

	1			5					10					15				
	Ser	Leu	Leu	Asn	Pro	Arg	His	Leu	Pro	Ser	Ile	Pro	Ala	Met	Phe	Pro		
5				20					25					30				
	Val	Ser	Ser	Gly	Cys	Phe	Gln	Glu	Gln	Gln	Glu	Met	Asn	Lys	Ser	Leu		
				35				40					45					
10	Val	Ser	Cys	Leu	Phe	Val	Leu	His	Phe	Val	Leu	His	Cys	Ile	Phe	Xaa		
			50				55					60						

15

(2) INFORMATION FOR SEQ ID NO: 324:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 196 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEO ID NO: 324:

25	Met Leu Ser Thr Ser Glu Tyr Ser Gln Ser Pro Lys Met Glu Ser Leu	
	1 5 10 15	
30	Ser Ser His Arg Ile Asp Glu Asp Gly Glu Asn Thr Gln Ile Glu Asp	
	20 25 30	
	Thr Glu Pro Met Ser Pro Val Leu Asn Ser Lys Phe Val Pro Ala Glu	
	35 40 45	
35	Asn Asp Ser Ile Leu Met Asn Pro Ala Gln Asp Gly Glu Val Gln Leu	
	50 55 60	
	Ser Gln Asn Asp Asp Lys Thr Lys Gly Asp Asp Thr Asp Thr Arg Asp	
	65 70 75 80	
40	Asp Ile Ser Ile Leu Ala Thr Gly Cys Lys Gly Arg Glu Glu Thr Val	
	85 90 95	
	Ala Glu Glu Val Cys Ile Asp Leu Thr Cys Asp Ser Gly Ser Gln Ala	
	100 105 110	
45	Val Pro Ser Pro Ala Thr Arg Ser Glu Ala Leu Ser Ser Val Leu Asp	
	115 120 125	
	Gln Glu Glu Ala Met Glu Ile Lys Glu His His Pro Glu Glu Gly Ser	
	130 135 140	
	Ser Gly Ser Glu Val Glu Ile Pro Glu Thr Pro Cys Glu Ser Gln	
	145 150 155 160	
55	Gly Glu Glu Leu Lys Glu Glu Asn Met Glu Ser Val Pro Leu His Leu	
	165 170 175	
	Ser Leu Thr Glu Thr Gln Ser Gln Gly Leu Cys Leu Arg Arg His Pro	
	180 185 190	
60		

Lys Lys Lys Lys
195

5

(2) INFORMATION FOR SEQ ID NO: 325:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 252 amino acids

10

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 325:

Met Gly Gly Asp Leu Val Leu Gly Leu Gly Ala Leu Arg Arg Arg Lys
1 5 10 15

Arg Leu Leu Glu Gln Glu Lys Ser Leu Ala Gly Trp Ala Leu Val Leu
20 25 30

Ala Xaa Xaa Gly Ile Gly Leu Met Val Leu His Ala Glu Met Leu Trp
35 40 45

Phe Gly Gly Cys Ser Ala Val Asn Ala Thr Gly His Leu Ser Asp Thr
50 55 60

Leu Trp Leu Ile Pro Ile Thr Phe Leu Thr Ile Gly Tyr Gly Asp Val
65 70 75 80

Val Pro Gly Thr Met Trp Gly Lys Ile Val Cys Leu Cys Thr Gly Val
85 90 95

Met Gly Val Cys Cys Thr Ala Leu Leu Val Ala Val Val Ala Arg Lys
100 105 110

Leu Glu Phe Asn Lys Ala Glu Lys His Val His Asn Phe Met Met Asp
115 120 125

Ile Gln Tyr Thr Lys Glu Met Lys Glu Ser Ala Ala Arg Val Leu Gln
130 135 140

Glu Ala Trp Met Phe Tyr Lys His Thr Arg Arg Lys Glu Ser His Ala
145 150 155 160

Ala Arg Xaa His Gln Arg Xaa Leu Leu Ala Ala Ile Asn Ala Phe Arg
165 170 175

Gln Val Arg Leu Lys His Arg Lys Leu Arg Glu Gln Val Asn Ser Met
180 185 190

Val Asp Ile Ser Lys Met His Met Ile Leu Tyr Asp Leu Gln Gln Asn
195 200 205

Leu Ser Ser Ser His Arg Ala Leu Glu Lys Gln Ile Asp Thr Leu Ala
210 215 220

Gly Lys Leu Asp Ala Leu Thr Glu Leu Leu Ser Thr Ala Leu Gly Pro
225 230 235 240

Arg Gln Leu Pro Glu Pro Ser Gln Gln Ser Lys Xaa
245 250

5 (2) INFORMATION FOR SEQ ID NO: 326:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 68 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 326:

Met Trp Arg Cys Arg Gly Lys Leu Ser Phe Pro Leu Phe Ala Val Val
1 5 10 15

15 Ile Val Ser Cys Arg Lys Asp Gly Pro Asp Ala Ala Ala Ala Pro Ala
20 25 30

Val Ile Lys Asn Asn Ser His Tyr Gln Thr Ser Lys Ala Leu Glu Leu
35 40 45

20 Glu Lys Thr Thr Glu Asn Lys Glu Ser Asn Pro Phe Ile Leu Gln Val
50 55 60

25 Asn Lys Leu Xaa
65

30 (2) INFORMATION FOR SEQ ID NO: 327:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 84 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 327:

Met Gly Glu Gly Lys Asn Gly Phe Gly Gly Phe Val His Thr Ala Asp
1 5 10 15

40 Ala Cys Trp Glu Gly Val His Ser Glu Pro Val Cys Arg Thr Val His
20 25 30

Thr Val His Thr Cys His His Gln Ala Phe Leu Val Leu Ile Gly Trp
35 40 45

45 Ser Lys Ser Gly Lys Glu Arg Lys Glu Ala Phe Leu Thr Ala Ile Ile
50 55 60

50 Leu Asn Ser Arg Ser Ile His Ile Ser Cys Ser Trp Pro Pro Ser Pro
65 70 75 80

Val Pro Gln Xaa

55

(2) INFORMATION FOR SEQ ID NO: 328:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 amino acids

60

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 328:

5 Met Leu Leu Ile Asn Leu Leu Trp Leu Val Thr Met Ile Lys Ser Val
1 5 10 15
Ile Asn Asn Asn Ile Ile Leu Phe Leu Lys Lys Lys Ser Leu Phe Phe
20 25 30
10 Ile Asp Ser Val
35

15

(2) INFORMATION FOR SEQ ID NO: 329:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 63 amino acids

20 (B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 329:

25 Met Thr Phe Pro Phe Glu Lys Lys Ile Val Ala Phe Ser Ala Phe Tyr
1 5 10 15
Leu Ile Pro Gly Glu Ser Arg Leu Ala Pro Thr Phe Asn Pro Ser Ala
20 25 30
30 Asp Met Thr Val Ile Leu Arg Gly Arg Ala Gln His Lys Thr Ala Met
35 40 45
Leu Glu Ser Tyr Asn Trp Lys Val Ser Cys Gln Leu Arg Glu Xaa
50 55 60
35

(2) INFORMATION FOR SEQ ID NO: 330:

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 35 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 330:

45

Met His Ser Lys Gly Ser Ser Leu Leu Leu Phe Leu Pro Gln Leu Ile
1 5 10 15
Leu Ile Leu Pro Val Cys Ala His Leu His Glu Glu Leu Asn Cys Cys
20 25 30
Phe His Arg
35

55

(2) INFORMATION FOR SEQ ID NO: 331:

(i) SEQUENCE CHARACTERISTICS:

60 (A) LENGTH: 23 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 331:

5 Met Gly Ala Leu Val Leu Leu Leu Cys Leu Leu Val Gly Val Gln Gln
 1 5 10 15

Ser Gly Ser Val Trp Asp Ser
 20

10

(2) INFORMATION FOR SEQ ID NO: 332:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 40 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 332:

20 Met Gln Ser Ala Glu Ile Leu Ser Trp Thr Asp Val Leu His Asp Phe
 1 5 10 15

Leu Phe Ser Leu Phe Leu Trp Pro Ala Phe Glu Asp Arg Ala Leu Leu
 25 20 25 30

Ile Phe Thr Leu Asn Gln Ile Val
 35 40

30

(2) INFORMATION FOR SEQ ID NO: 333:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 111 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 333:

40 Met Gln Ser Leu Val Gln Trp Gly Leu Asp Ser Tyr Asp Tyr Leu Gln
 1 5 10 15

Asn Ala Pro Pro Gly Phe Phe Pro Arg Leu Gly Val Ile Gly Phe Ala
 20 25 30

45 Gly Leu Ile Gly Leu Leu Leu Ala Arg Gly Ser Lys Ile Lys Lys Leu
 35 40 45

Val Tyr Pro Pro Gly Phe Met Gly Leu Ala Ala Ser Leu Tyr Tyr Pro
 50 50 55 60

Gln Gln Ala Ile Val Phe Ala Gln Val Ser Gly Glu Arg Leu Tyr Asp
 65 70 75 80

55 Trp Gly Leu Arg Gly Tyr Ile Val Ile Glu Asp Leu Trp Lys Glu Asn
 85 90 95

Phe Gln Lys Pro Gly Asn Val Lys Asn Ser Pro Gly Thr Lys Xaa
 100 105 110

60

(2) INFORMATION FOR SEQ ID NO: 334:

- 5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 106 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 334:

10 Met Ala Pro Ser Leu Leu Leu Leu Ala Pro Leu Cys Ser Leu Glu Ala
 1 5 10 15
 15 Val Leu Ser Ser Pro Leu Glu Lys Gln Cys Gln Leu Pro Gly Ile Phe
 20 25 30
 Cys Gln Leu Gln Leu Pro Cys Pro Leu Leu Leu Ser Ala Gln Leu Leu
 35 40 45
 20 Lys Gly Ile Val Xaa Pro Arg Cys Pro Ala Ser Leu Pro Gln Pro Pro
 50 55 60
 His Pro Ala Pro Ser Trp His Leu Pro Leu His Cys Thr Glu Arg Xaa
 65 70 75 80
 25 Pro His His Leu Pro Leu Gln Gly Gly Ser Ser Asn Met Glu Glu Xaa
 85 90 95
 30 Asn Tyr Arg Gly Tyr Xaa Asp Ala Gln Leu
 100 105

(2) INFORMATION FOR SEQ ID NO: 335:

- 35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 50 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 335:

Met Thr Thr Cys Leu Phe Gly Leu Leu Ser Cys Glu Met Ser Ala Gln
 1 5 10 15
 45 Val Ser Gln Lys Ser Cys Val Tyr Asp Glu Ser Glu Cys Phe Ser Ser
 20 25 30
 Val Gly Gln Leu Leu Ala Leu Leu Ile Leu Val Tyr Val Leu Pro Ser
 35 40 45
 50 Ile Xaa
 50

55

(2) INFORMATION FOR SEQ ID NO: 336:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 48 amino acids
 (B) TYPE: amino acid
- 60

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 336:

5 Met Leu Trp Lys Cys Ser Gln Asn Ile Ala Arg Cys Leu Leu Leu
 1 5 10 15

Leu Ala Leu Val Glu Ile Lys Leu Glu Asp Leu Gln Ser Gln Leu His
 20 25 30

10 Pro Thr Trp Lys Ser Ile Pro Gly Pro Ser Pro Arg Asn Gln His Arg
 35 40 45

15

(2) INFORMATION FOR SEQ ID NO: 337:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 41 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 337:

25 Met Leu Ile Pro Leu Gln Cys Leu Phe Ser Ser Asp Arg Met Leu Thr
 1 5 10 15

30 Phe Leu Thr Pro Trp Gln Lys Gly Glu Lys Cys Val Leu Gly Trp Val
 20 25 30

Thr Lys Phe Leu Ser Glu Ile Ser Xaa
 35 40

35

(2) INFORMATION FOR SEQ ID NO: 338:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 76 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 338:

45 Met Thr Phe Ser Ser Leu Lys Leu Phe Val Leu Thr Cys Ile Ile Lys
 1 5 10 15

Gly Leu Glu Arg Phe Ile Ile Leu Arg Glu Val Cys Asn Gln Glu Ile
 20 25 30

50 Gln Arg Ser Leu Ser Ser Asn Leu Val His Val Leu Leu Gln Pro Ala
 35 40 45

55 Thr Phe Lys Asp Val Leu Val Thr Glu Ile Ile Cys Leu Cys Met Cys
 50 55 60

Leu Tyr Ser Ile Lys Tyr Met Pro Pro Gln Lys Lys
 65 70 75

60

(2) INFORMATION FOR SEQ ID NO: 339:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 339:

5
10 Lys Val Tyr Ile Phe Leu Ile Phe Met Val Leu Ile Leu Pro Ser Leu
1 5 10 15
Gly Leu Thr Arg Tyr Met Pro Pro Xaa Ser Xaa Leu Asn Ser Glu
20 25 30
15

(2) INFORMATION FOR SEQ ID NO: 340:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 42 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 340:

20
25 Met Ala Lys Ile Ser Pro Phe Glu Val Val Lys Arg Thr Ser Val Pro
1 5 10 15
30 Val Leu Val Gly Leu Val Ile Val Ile Val Ala Thr Glu Leu Met Val
20 25 30
Pro Gly Thr Ala Ala Ala Val Thr Gly Lys
35 40
35

(2) INFORMATION FOR SEQ ID NO: 341:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 341:

40
45 Met Arg Leu Phe Phe Ile Gly Phe Leu Leu Leu Phe Ser Phe Gly Leu
1 5 10 15
Leu Arg Gln Pro Ser Leu Ser Ala Glu His
20 25
50

(2) INFORMATION FOR SEQ ID NO: 342:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 342:

60

Met Val Phe Ser Val Ser Ser Ala Leu Ala Leu Leu Leu Met Leu Leu
 1 5 10 15

Arg Ser Ser Asp Leu Ala Lys Lys Thr Glu
 20 25

(2) INFORMATION FOR SEQ ID NO: 343:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 157 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 343:

Met Ser Leu Glu Phe Tyr Gln Lys Lys Lys Ser Arg Trp Pro Phe Ser
 1 5 10 15

Asp Glu Cys Ile Pro Trp Glu Val Trp Thr Val Lys Val His Val Val
 20 25 30

Ala Leu Ala Thr Glu Gln Glu Arg Gln Ile Cys Arg Glu Lys Val Gly
 35 40 45

Glu Lys Leu Cys Glu Lys Ile Ile Asn Ile Val Glu Val Met Asn Arg
 50 55 60

His Glu Tyr Leu Pro Lys Met Pro Thr Gln Ser Glu Val Asp Asn Val
 65 70 75 80

Phe Asp Thr Gly Leu Arg Asp Val Gln Pro Tyr Leu Tyr Lys Ile Ser
 85 90 95

Phe Gln Ile Thr Asp Ala Leu Gly Thr Ser Val Thr Thr Thr Met Arg
 100 105 110

Arg Leu Ile Lys Asp Thr Leu Pro Ser Glu Arg Arg Trp Ile Ser Gly
 115 120 125

Ser Ser Leu Met Ala Pro Arg Pro Trp Leu Leu Gly Ile Ala Leu Leu
 130 135 140

Gly Leu Trp Ala Leu Glu Pro Ala Leu Gly His Trp Xaa
 145 150 155

(2) INFORMATION FOR SEQ ID NO: 344:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 520 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 344:

Met Phe Leu Leu Pro Leu Pro Ala Ala Gly Arg Val Val Val Arg Arg
 1 5 10 15

Leu Ala Val Arg Arg Phe Gly Ser Arg Ser Leu Ser Thr Ala Asp Met

	20	25	30
	Thr Lys Gly Leu Val Leu Gly Ile Tyr Ser Lys Glu Lys Glu Asp Asp		
	35	40	45
5	Val Pro Gln Phe Thr Ser Ala Gly Glu Asn Phe Asp Lys Leu Leu Ala		
	50	55	60
10	Gly Lys Leu Arg Glu Thr Leu Asn Ile Ser Gly Pro Pro Leu Lys Ala		
	65	70	75
	Gly Lys Thr Arg Thr Phe Tyr Gly Leu His Gln Asp Phe Pro Ser Val		
	85	90	95
15	Val Leu Val Gly Leu Gly Lys Lys Ala Ala Gly Ile Asp Glu Gln Glu		
	100	105	110
	Asn Trp His Glu Gly Lys Glu Asn Ile Arg Ala Ala Val Ala Ala Gly		
	115	120	125
20	Cys Arg Gln Ile Gln Asp Leu Glu Leu Ser Ser Val Glu Val Asp Pro		
	130	135	140
	Cys Gly Asp Ala Gln Ala Ala Glu Gly Ala Val Leu Gly Leu Tyr		
	145	150	155
25	Glu Tyr Asp Asp Leu Lys Gln Lys Lys Lys Met Ala Val Ser Ala Lys		
	165	170	175
30	Leu Tyr Gly Ser Gly Asp Gln Glu Ala Trp Gln Lys Gly Val Leu Phe		
	180	185	190
	Ala Ser Gly Gln Asn Leu Ala Arg Gln Leu Met Glu Thr Pro Ala Asn		
	195	200	205
35	Glu Met Thr Pro Thr Arg Phe Ala Glu Ile Ile Glu Lys Asn Leu Lys		
	210	215	220
	Ser Ala Ser Ser Lys Thr Glu Val His Ile Arg Pro Lys Ser Trp Ile		
	225	230	235
	Glu Glu Gln Ala Met Gly Ser Phe Leu Ser Val Ala Lys Gly Ser Asp		
	245	250	255
45	Glu Pro Pro Val Phe Leu Glu Ile His Tyr Lys Gly Ser Pro Asn Ala		
	260	265	270
	Asn Glu Pro Pro Leu Val Phe Val Gly Lys Gly Ile Thr Phe Asp Ser		
	275	280	285
50	Gly Gly Ile Ser Ile Lys Ala Ser Ala Asn Met Asp Leu Met Arg Ala		
	290	295	300
	Asp Met Gly Gly Ala Ala Thr Ile Cys Ser Ala Ile Val Ser Ala Ala		
	305	310	315
55	Lys Leu Asn Leu Pro Ile Asn Ile Ile Gly Leu Ala Pro Leu Cys Glu		
	325	330	335
60	Asn Met Pro Ser Gly Lys Ala Asn Lys Pro Gly Asp Val Val Arg Ala		

	340	345	350
	Lys Asn Gly Lys Thr Ile Gln Val Asp Asn Thr Asp Ala Glu Gly Arg		
	355	360	365
5	Leu Ile Leu Ala Asp Ala Leu Cys Tyr Ala His Thr Phe Asn Pro Lys		
	370	375	380
10	Xaa Ile Leu Asn Ala Ala Thr Leu Thr Gly Ala Met Asp Val Ala Leu		
	385	390	395
	Gly Ser Gly Ala Thr Gly Val Phe Thr Asn Ser Ser Trp Leu Trp Asn		
	405	410	415
15	Lys Leu Phe Glu Ala Ser Ile Glu Thr Gly Asp Arg Val Trp Arg Met		
	420	425	430
	Pro Leu Phe Glu His Tyr Thr Arg Gln Val Val Asp Cys Gln Leu Ala		
	435	440	445
20	Asp Val Asn Asn Ile Gly Lys Tyr Arg Ser Ala Gly Ala Cys Thr Ala		
	450	455	460
	Ala Ala Phe Leu Lys Glu Phe Val Thr His Pro Lys Trp Ala His Leu		
25	465	470	475
	Asp Ile Ala Gly Val Met Thr Asn Lys Asp Glu Val Pro Tyr Leu Arg		
	485	490	495
30	Lys Gly Met Thr Gly Arg Pro Thr Arg Thr Leu Ile Glu Phe Leu Leu		
	500	505	510
	Arg Phe Ser Gln Asp Asn Ala Xaa		
	515	520	

35

(2) INFORMATION FOR SEQ ID NO: 345:

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 345:

45

Thr	Ile	Leu	Phe	Leu	Phe	Leu	Gln	Leu	Ser	Ala	Leu	Arg	Leu	Ile	Val
1				5					10					15	

50

Gly	Lys	Asp	Ser	Ile	Asp	Ile	Asp	Ile	Ser	Ser	Arg	Arg	Arg	Glu	Asp
			20					25						30	

Gln	Ser	Leu	Arg	Leu	Asn	Ala
				35		

55

(2) INFORMATION FOR SEQ ID NO: 346:

60

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 234 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 346:

5 Met Thr Ser Glu Leu Asp Ile Phe Val Gly Asn Thr Thr Leu Ile Asp
 1 5 10 15
 Glu Asp Val Tyr Arg Leu Trp Leu Asp Gly Tyr Ser Val Thr Asp Ala
 20 25 30
 10 Val Ala Leu Arg Val Arg Ser Gly Ile Leu Glu Gln Thr Gly Ala Thr
 35 40 45
 Ala Ala Val Leu Gln Ser Asp Thr Met Asp His Tyr Arg Thr Phe His
 15 50 55 60
 Met Leu Glu Arg Leu Leu His Ala Pro Pro Lys Leu Leu His Gln Leu
 65 70 75 80
 20 Ile Phe Gln Ile Pro Pro Ser Arg Gln Ala Leu Leu Ile Glu Arg Tyr
 85 90 95
 Tyr Ala Phe Asp Glu Ala Phe Val Arg Glu Val Leu Gly Lys Lys Leu
 100 105 110
 25 Ser Lys Gly Thr Lys Lys Asp Leu Asp Asp Ile Ser Thr Lys Thr Gly
 115 120 125
 Ile Thr Leu Lys Ser Cys Arg Arg Gln Phe Asp Asn Phe Lys Arg Val
 30 130 135 140
 Phe Lys Val Val Glu Glu Met Arg Gly Ser Leu Val Asp Asn Ile Gln
 145 150 155 160
 35 Gln His Phe Leu Leu Ser Asp Arg Leu Ala Arg Asp Tyr Ala Ala Ile
 165 170 175
 Val Phe Phe Ala Asn Asn Arg Phe Glu Thr Gly Lys Lys Lys Leu Gln
 180 185 190
 40 Tyr Leu Ser Phe Gly Asp Phe Ala Phe Cys Ala Glu Leu Met Ile Gln
 195 200 205
 Asn Trp Thr Leu Gly Pro Val Asp Ser Gln Met Asp Asp Met Asp Met
 45 210 215 220
 Asp Leu Asp Arg Asn Phe Ser Arg Thr Xaa
 225 230
 50

(2) INFORMATION FOR SEQ ID NO: 347:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 169 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 347:

60 Met Ala Ala Ala Val Ala Gly Met Leu Arg Gly Gly Leu Leu Pro Gln

1 5 10 15
 Ala Gly Arg Leu Pro Thr Leu Gln Thr Val Arg Tyr Gly Ser Lys Ala
 20 25 30
 5 Val Thr Arg His Arg Arg Val Met His Phe Gln Arg Gln Lys Leu Met
 35 40 45
 10 Ala Val Thr Glu Tyr Ile Pro Pro Lys Pro Ala Ile His Pro Ser Cys
 50 55 60
 Leu Pro Ser Pro Pro Ser Pro Pro Gln Glu Glu Ile Gly Leu Ile Arg
 65 70 75 80
 15 Leu Leu Arg Arg Glu Ile Ala Ala Val Phe Gln Asp Asn Arg Met Ile
 85 90 95
 Ala Val Cys Gln Asn Val Ala Leu Ser Ala Glu Asp Lys Leu Leu Ile
 100 105 110
 20 Ala Thr Pro Ala Ala Glu Thr Gln Asp Pro Asp Glu Gly Leu Pro Gln
 115 120 125
 25 Pro Gly Pro Glu Ser Pro Ser Trp Arg Ile Pro Ser Thr Lys Ile Cys
 130 135 140
 Cys Pro Phe Leu Trp Gly Thr Thr Cys Cys Trp Ser Val Lys Ser Pro
 145 150 155 160
 30 Arg Ser Arg Arg Trp Tyr Gly Ser Xaa
 165

35 (2) INFORMATION FOR SEQ ID NO: 348:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 40 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 348:

Met Lys Arg Ser Phe Leu Leu Pro Leu Leu Leu Val Gly Phe Leu Asp
 1 5 10 15
 45 Thr Ala His Leu Ile Leu Leu Glu Thr Leu Ser Val Cys Leu Trp Leu
 20 25 30
 50 Pro Ser Leu Ile Asp Ser Arg Cys Val Met Ser
 35 40

55 (2) INFORMATION FOR SEQ ID NO: 349:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 78 amino acids
 (B) TYPE: amino acid
 60 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 349:

Met Lys Glu Gly Pro Pro Cys Lys Arg His His Tyr Tyr Gln Asn Cys
 1 5 10 15

5 Gly Ala Lys Leu Val Ser Leu Phe Gly Glu Thr Asn Gln Ile His
 20 25 30

Leu Leu Glu Thr Gln Val Gly Thr Glu Lys Gly Gly Glu Arg Ile Trp
 35 40 45

10 Glu Glu Lys Trp Arg Ile Ser Ser Thr Val Leu Phe Ile Ser Val Asn
 50 55 60

15 Ser Tyr Val Glu Gly Ser Val Leu Glu Ile Lys Leu Phe Tyr
 65 70 75

20 (2) INFORMATION FOR SEQ ID NO: 350:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 350:

Met Ser Glu Ile Leu Ser Leu Leu Phe Cys Leu Leu Gly Pro Ala Leu
 1 5 10 15

30 Asp Glu Arg Arg Glu Glu Lys Asp
 20

35 (2) INFORMATION FOR SEQ ID NO: 351:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 274 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 351:

Met Ser Ser Ala Gly Thr Ala Thr Pro Leu Glu Met Asp His Lys Leu
 1 5 10 15

45 Thr Ser Gln Pro Gly Arg Pro Ser Phe Tyr Cys Asn Ser Arg His Ser
 20 25 30

50 Ile Val Gly Ser Ser His Gln Leu Gly Phe Trp Phe Ser His Leu Glu
 35 40 45

Ser Ser Gly Leu Lys Val Phe Gln Val Ser Leu Pro Cys Glu Cys Val
 50 55 60

55 Asn Leu Pro Thr Arg Ile Ala Ser Val Val Leu Ser Leu Met Ser Leu
 65 70 75 80

60 Leu Val Val Gly Gln Ala Pro Ala Trp Glu Gly Ser Leu Leu Arg Gly
 85 90 95

Arg Pro Ala Gly Gly Ala His Leu Cys Ala Met Xaa Val Ile Glu Gly
 100 105 110
 5 Leu Val Val Asp Val Gly Glu Arg Ile Leu His Gly Gln Arg Glu Val
 115 120 125
 Gly Gln Val Ser Gln Val Leu Pro Ala Leu Ser Leu Gly Leu Val Phe
 130 135 140
 10 Leu Cys Gln Gly Thr Val Glu Lys Val Ser Gly Ala Ala His Cys Ser
 145 150 155 160
 Ser Leu Leu Cys Cys Leu Pro Trp Gln Cys Ser Gly Gly Gly Phe Pro
 165 170 175
 15 Thr Xaa Arg Cys Ser Arg Pro Tyr Phe Ser Ser His Lys Gly Val Ala
 180 185 190
 20 Ala Thr Leu Ala Leu Thr Cys His Cys Asp Lys Val His Val Ala Gly
 195 200 205
 Leu Gly Lys Asp Trp Ala Ile Glu Gln Arg Arg Arg Thr Cys Glu Ser
 210 215 220
 25 Asp Xaa Glu Xaa Xaa Pro Phe Thr Leu Ala Gly Leu Val Leu Val Leu
 225 230 235 240
 Arg Phe Cys Gln Val Val Leu Val Trp Ile Pro Gln Leu Gly Asp Lys
 245 250 255
 30 His Trp Arg Gly Met Thr Arg Leu Gly Arg Val Ser Leu Thr Ser Ser
 260 265 270
 Ile Xaa

35

40 (2) INFORMATION FOR SEQ ID NO: 352:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 352:

Met Ile Phe Thr Ser Val Thr Lys Gly Ile Leu Leu Ile Ala Leu Trp
 1 5 10 15
 50 Val Pro Leu Phe His Phe Met Leu Ile Asp Ser Ile Leu Gly Pro Ser
 20 25 30
 Arg Leu Leu Thr Asp Gly Val Pro Phe Asn Pro Trp His Val Xaa
 35 40 45
 55

(2) INFORMATION FOR SEQ ID NO: 353:

60 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 353:

5

Met Lys Thr

1

10

(2) INFORMATION FOR SEQ ID NO: 354:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 52 amino acids

15

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 354:

20

Met Ser Ile Ser Gly Thr Asp Gly Leu Ile Leu Leu Val Gly Leu
1 5 10 15Glu Ala Xaa Val Arg Ser Ser Lys Lys Trp Ile Pro Lys Ala Leu Xaa
20 25 30

25

Val Thr Gln Ala Lys Trp Asn Ser Trp Pro Ser Arg Arg Asn Ala Gly
35 40 45Phe Ala Leu His
50

30

(2) INFORMATION FOR SEQ ID NO: 355:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 132 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 355:

40

Met Glu His Cys Leu Tyr His Ser Val His Gly Ile Asn Pro Tyr Ile
1 5 10 15

45

His Lys Asn Thr His Pro Ser Ile Asn Ile Tyr Met Val Trp Asp Glu
20 25 30Gln Val Asn Ser Phe Glu Arg Glu Phe Val Pro Phe Phe Phe Leu Ile
35 40 45

50

Ile Leu Leu Asn Cys Cys Gln Leu Ser Asn Lys Gln Thr Glu Lys Leu
50 55 60Phe Gly Lys Thr Leu His Thr Pro Phe Leu Ser Ser Ala Leu Lys Tyr
65 70 75 80

55

Arg Leu Asn Thr His Ile Leu Pro Val Phe Ser Tyr Ser Asp Ser Ile
85 90 95

60

Leu Thr Cys His Leu Ile Leu Ala Ser Tyr Phe Ser His Val Tyr Leu
100 105 110

Pro Val Thr Cys Ile Cys Tyr Leu Asn Arg Lys Lys Asn Ile Gln Lys
 115 120 125

5 Lys Lys Asn Xaa
 130

10 (2) INFORMATION FOR SEQ ID NO: 356:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 204 amino acids

(B) TYPE: amino acid

15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 356:

Met Gly Ser Arg Asp His Leu Phe Lys Val Leu Val Val Gly Asp Ala
 1 5 10 15
 20 Ala Val Gly Lys Thr Ser Leu Val Gln Asp Tyr Ser Gln Asp Ser Phe
 20 25 30
 Ser Lys His Tyr Lys Ser Thr Val Gly Val Asp Phe Ala Leu Lys Val
 25 35 40 45
 Leu Gln Trp Ser Asp Tyr Glu Ile Val Arg Leu Gln Leu Trp Asp Ile
 50 55 60
 30 Ala Gly Gln Glu Arg Phe Thr Ser Met Thr Arg Leu Tyr Tyr Arg Asp
 65 70 75 80
 Ala Ser Ala Cys Val Ile Met Phe Asp Val Thr Asn Ala Thr Thr Phe
 85 90 95
 35 Ser Asn Ser Gln Arg Trp Lys Gln Asp Leu Asp Ser Lys Leu Thr Leu
 100 105 110
 Pro Asn Gly Glu Pro Val Pro Cys Leu Leu Leu Ala Asn Lys Cys Asp
 40 115 120 125
 Leu Ser Pro Trp Ala Val Ser Arg Asp Gln Ile Asp Arg Phe Ser Lys
 130 135 140
 45 Glu Asn Gly Phe Thr Gly Trp Thr Glu Thr Ser Val Lys Glu Asn Lys
 145 150 155 160
 Asn Ile Asn Glu Ala Met Arg Val Leu Ile Glu Lys Met Met Arg Asn
 165 170 175
 50 Ser Thr Glu Asp Ile Met Ser Leu Ser Thr Gln Gly Asp Tyr Ile Asn
 180 185 190
 Leu Gln Thr Lys Ser Ser Ser Trp Ser Cys Cys Xaa
 55 195 200

60 (2) INFORMATION FOR SEQ ID NO: 357:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 357:

Met Ile Ser Leu Ile Phe Gln Leu Glu Glu Glu Lys Leu Val Glu Lys
 1 5 10 15
 10 Phe Phe Phe Phe Leu Phe Phe Phe Leu Lys Lys Gly Ser Gln Gly Ser
 20 25 30
 Asn Leu Lys Ile Val Pro Arg His Met Arg Val Val Leu Arg Gly
 35 40 45
 15

(2) INFORMATION FOR SEQ ID NO: 358:

20 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 73 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 358:

Met Thr Tyr Val Thr Cys Leu His Val Cys Leu Leu Val Glu Phe Leu
 1 5 10 15
 30 Asn Ser Gln Leu Thr Asn His Arg Lys Tyr Tyr Phe Leu Ser Tyr Gly
 20 25 30
 Phe Trp Phe Thr Gly Leu Arg Gly Phe Ser Glu Tyr Leu Trp Pro Gln
 35 40 45
 35 Gln His Thr Ser Phe His Pro Asn Arg Asn Glu Ile Asn Phe Val Ser
 50 55 60
 Thr Asp Asn Arg Ile Trp Val Thr Xaa
 65 70
 40

(2) INFORMATION FOR SEQ ID NO: 359:

45 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 102 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 359:

Met Ser Asp Gln Glu Ala Lys Pro Ser Thr Glu Asp Leu Gly Asp Lys
 1 5 10 15
 55 Lys Glu Gly Glu Tyr Ile Lys Leu Lys Val Ile Gly Gln Asp Ser Ser
 20 25 30
 Glu Ile His Phe Lys Val Lys Met Thr Thr His Leu Lys Lys Leu Lys
 35 40 45
 60 Glu Ser Tyr Cys Gln Arg Gln Gly Val Pro Met Asn Ser Leu Arg Phe

50 55 60

Leu Phe Glu Gly Gln Arg Ile Ala Asp Asn His Thr Pro Lys Glu Leu
65 70 75 80

Gly Met Glu Glu Glu Asp Val Ile Glu Val Tyr Gln Glu Gln Thr Gly
85 90 95

Gly His Ser Thr Val Xaa
100

(2) INFORMATION FOR SEQ ID NO: 360:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 48 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 360:

	Met	Gly	Phe	Pro	Gln	Trp	His	Leu	Gly	Asn	His	Ala	Val	Glu	Pro	Val
	1				5					10					15	
25	Thr	Ser	Ile	Leu	Leu	Leu	Phe	Leu	Leu	Met	Met	Leu	Gly	Val	Arg	Gly
				20					25					30		
	Leu	Leu	Leu	Val	Gly	Leu	Val	Tyr	Leu	Val	Ser	His	Leu	Ser	Gln	Arg
30			35					40					45			

35 (2) INFORMATION FOR SEQ ID NO: 361:

(i) SEQUENCE CHARACTERISTICS:

40 (A) LENGTH: 179 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 361:

45	Met	Ser	Ala	Glu	Val	Lys	Val	Thr	Gly	Gln	Asn	Gln	Glu	Gln	Phe	Leu
	1				5					10					15	
	Leu	Leu	Ala	Lys	Ser	Ala	Lys	Gly	Ala	Ala	Leu	Ala	Thr	Leu	Ile	His
				20					25					30		
50	Gln	Val	Leu	Glu	Ala	Pro	Gly	Val	Tyr	Val	Phe	Gly	Glu	Leu	Leu	Asp
			35					40					45			
	Met	Pro	Asn	Val	Arg	Glu	Leu	Ala	Glu	Ser	Asp	Phe	Ala	Ser	Thr	Phe
		50					55					60				
55	Arg	Leu	Leu	Thr	Val	Phe	Ala	Tyr	Gly	Thr	Tyr	Ala	Asp	Tyr	Leu	Ala
	65					70					75					80
	Glu	Ala	Arg	Asn	Leu	Pro	Pro	Leu	Thr	Glu	Ala	Gln	Lys	Asn	Lys	Leu
60					85					90					95	

Arg His Leu Ser Val Val Thr Leu Ala Ala Lys Val Lys Cys Ile Pro
 100 105 110

5 Tyr Ala Val Leu Leu Glu Ala Leu Ala Leu Arg Asn Val Arg Gln Leu
 115 120 125

Glu Asp Leu Val Ile Glu Ala Val Tyr Ala Asp Val Leu Arg Gly Ser
 130 135 140

10 Leu Asp Gln Arg Asn Gln Arg Leu Glu Val Asp Tyr Ser Ile Gly Arg
 145 150 155 160

15 Asp Ile Gln Arg Gln Asp Leu Ser Ala Ile Ala Arg Thr Leu Xaa Lys
 165 170 175

Asn His Xaa

20

(2) INFORMATION FOR SEQ ID NO: 362:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 362:

30 Met Lys Ser Ser Ser Leu Phe Phe Phe Phe Leu Ala His Phe Ile His
 1 5 10 15

Ser His Asp Leu Pro Gly Leu Cys Arg
 20 25

35

(2) INFORMATION FOR SEQ ID NO: 363:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 224 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 363:

45 Met Lys Phe Ala Ala Ser Gly Xaa Phe Leu His His Met Ala Gly Leu
 1 5 10 15

50 Ser Ser Ser Lys Leu Ser Met Ser Lys Ala Leu Pro Leu Thr Lys Val
 20 25 30

Val Gln Asn Asp Ala Tyr Thr Ala Pro Ala Leu Pro Ser Ser Ile Arg
 35 40 45

55 Thr Lys Ala Leu Thr Asn Met Ser Arg Thr Leu Val Asn Lys Glu Glu
 50 55 60

Pro Pro Lys Glu Leu Pro Ala Ala Glu Pro Val Leu Ser Pro Leu Glu
 65 70 75 80

60

	Gly	Thr	Lys	Met	Thr	Val	Asn	Asn	Leu	His	Pro	Arg	Val	Thr	Glu	Glu
					85				90						95	
5	Asp	Ile	Val	Glu	Leu	Phe	Cys	Val	Cys	Gly	Ala	Leu	Lys	Arg	Ala	Arg
				100					105					110		
	Leu	Val	His	Pro	Gly	Val	Ala	Glu	Val	Val	Phe	Val	Lys	Lys	Asp	Asp
				115				120					125			
10	Ala	Ile	Thr	Ala	Tyr	Lys	Lys	Tyr	Asn	Asn	Arg	Cys	Leu	Asp	Gly	Gln
		130					135					140				
	Pro	Met	Lys	Cys	Asn	Leu	His	Met	Asn	Gly	Asn	Val	Ile	Thr	Ser	Asp
15	145					150					155					160
	Gln	Pro	Ile	Leu	Leu	Arg	Leu	Ser	Asp	Ser	Pro	Ser	Met	Lys	Lys	Glu
					165					170					175	
	Ser	Glu	Leu	Pro	Arg	Arg	Val	Asn	Ser	Ala	Ser	Ser	Ser	Asn	Pro	Pro
20				180					185					190		
	Ala	Glu	Val	Asp	Pro	Asp	Thr	Ile	Leu	Lys	Ala	Leu	Phe	Lys	Ser	Ser
			195					200					205			
25	Gly	Ala	Ser	Xaa	Thr	Thr	Gln	Pro	Thr	Glu	Phe	Lys	Ile	Lys	Leu	Xaa
	210						215					220				

30

(2) INFORMATION FOR SEQ ID NO: 364:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 349 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 364:

```

40      Met Ser Lys Asn Cys Ile Lys Leu Leu Cys Glu Asp Pro Val Phe Ala
          1              5              10              15

45      Glu Tyr Ile Lys Cys Ile Leu Met Asp Glu Arg Thr Phe Leu Asn Asn
          20              25              30

      Asn Ile Val Tyr Thr Phe Met Thr His Phe Leu Leu Lys Val Gln Ser
          35              40              45

50      Gln Val Phe Ser Glu Ala Asn Cys Ala Asn Leu Ile Ser Thr Leu Ile
          50              55              60

      Thr Asn Leu Ile Ser Gln Tyr Gln Asn Leu Gln Ser Asp Phe Ser Asn
          65              70              75              80

55      Arg Val Glu Ile Ser Lys Ala Ser Ala Ser Leu Asn Gly Asp Leu Arg
          85              90              95

      Ala Leu Ala Leu Leu Leu Ser Val His Thr Pro Lys Gln Leu Asn Pro
          100              105              110

```

Ala Leu Ile Pro Thr Leu Gln Glu Leu Leu Ser Lys Cys Arg Thr Cys
115 120 125

5 Leu Gln Gln Arg Asn Ser Leu Gln Glu Gln Glu Ala Lys Glu Arg Lys
130 135 140

Thr Lys Asp Asp Glu Gly Ala Thr Pro Ile Lys Arg Arg Arg Val Ser
145 150 155 160

10 Ser Asp Glu Glu His Thr Val Asp Ser Cys Ile Ser Asp Met Lys Thr
165 170 175

Glu Thr Arg Glu Val Leu Thr Pro Thr Ser Thr Ser Asp Asn Glu Thr
15 180 185 190

Arg Asp Ser Ser Ile Ile Asp Pro Gly Thr Glu Gln Asp Leu Pro Ser
195 200 205

20 Pro Glu Asn Ser Ser Val Lys Glu Tyr Arg Met Glu Val Pro Ser Ser
210 215 220

Phe Ser Glu Asp Met Ser Asn Ile Arg Ser Gln His Ala Glu Glu Gln
225 230 235 240

25 Ser Asn Asn Gly Arg Tyr Asp Asp Cys Lys Glu Phe Lys Asp Leu His
245 250 255

Cys Ser Lys Asp Ser Thr Leu Ala Glu Glu Ser Glu Phe Pro Ser
30 260 265 270

Thr Ser Ile Ser Ala Val Leu Ser Asp Leu Ala Asp Leu Arg Ser Cys
275 280 285

35 Asp Gly Gln Ala Leu Pro Ser Gln Asp Pro Glu Val Ala Leu Ser Leu
290 295 300

Ser Cys Gly His Ser Arg Gly Leu Phe Ser His Met Gln Gln His Asp
40 305 310 315 320

Ile Leu Asp Thr Leu Cys Arg Thr Ile Glu Ser Thr Ile His Val Val
325 330 335

45 Thr Arg Ile Ser Gly Lys Gly Asn Gln Ala Ala Ser Xaa
340 345

(2) INFORMATION FOR SEQ ID NO: 365:

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 467 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 365:

Met Leu His Gln Asp His Ile Thr Phe Ala Met Leu Leu Ala Arg Ile
1 5 10 15

60 Lys Leu Lys Gly Thr Val Gly Glu Pro Thr Tyr Asp Ala Glu Phe Gln

	20	25	30
	His Phe Leu Arg Gly Asn Glu Ile Val Leu Ser Ala Gly Ser Thr Pro		
	35	40	45
5	Arg Ile Gln Gly Leu Thr Val Glu Gln Ala Glu Ala Val Val Arg Leu		
	50	55	60
10	Ser Cys Leu Pro Ala Phe Lys Asp Leu Ile Ala Lys Val Gln Ala Asp		
	65	70	75
	Glu Gln Phe Gly Ile Trp Leu Asp Ser Ser Ser Pro Glu Gln Thr Val		
	85	90	95
15	Pro Tyr Leu Trp Ser Glu Glu Thr Pro Ala Thr Pro Ile Gly Gln Ala		
	100	105	110
	Ile His Arg Leu Leu Leu Ile Gln Ala Phe Arg Pro Asp Arg Leu Leu		
	115	120	125
20	Ala Met Ala His Met Phe Val Ser Thr Asn Leu Gly Glu Ser Phe Met		
	130	135	140
	Ser Ile Met Glu Gln Pro Leu Asp Leu Thr His Ile Val Xaa Thr Glu		
	145	150	155
25	Val Lys Pro Asn Thr Pro Val Leu Met Cys Ser Val Pro Gly Tyr Asp		
	165	170	175
30	Ala Ser Gly His Val Glu Asp Leu Ala Ala Glu Gln Asn Thr Gln Ile		
	180	185	190
	Thr Ser Ile Ala Ile Gly Ser Ala Glu Gly Phe Asn Gln Ala Asp Lys		
	195	200	205
35	Ala Ile Asn Thr Ala Val Lys Ser Gly Arg Trp Val Met Leu Lys Asn		
	210	215	220
	Val His Leu Ala Pro Gly Trp Leu Met Gln Leu Glu Lys Lys Leu His		
	225	230	235
40	Ser Leu Gln Pro His Ala Cys Phe Arg Leu Phe Leu Thr Met Glu Ile		
	245	250	255
45	Asn Pro Lys Val Pro Val Asn Leu Leu Arg Ala Gly Arg Ile Phe Val		
	260	265	270
	Phe Glu Pro Pro Pro Gly Xaa Lys Ala Asn Met Leu Arg Thr Phe Ser		
	275	280	285
50	Ser Ile Pro Val Ser Arg Ile Cys Lys Ser Pro Asn Glu Arg Ala Arg		
	290	295	300
	Leu Tyr Phe Leu Leu Ala Trp Phe His Ala Ile Ile Gln Glu Arg Leu		
	305	310	315
55	Arg Tyr Ala Pro Leu Gly Trp Ser Lys Lys Tyr Glu Phe Gly Glu Ser		
	325	330	335
60	Asp Leu Arg Ser Xaa Cys Asp Thr Val Asp Thr Trp Leu Asp Asp Thr		

340 345 350
 Ala Lys Gly Arg Gln Asn Ile Ser Pro Asp Lys Ile Pro Trp Ser Ala
 355 360 365
 5 Leu Lys Thr Leu Met Ala Gln Ser Ile Tyr Gly Gly Arg Val Asp Asn
 370 375 380
 10 Glu Phe Asp Gln Arg Leu Leu Asn Thr Phe Leu Glu Arg Leu Phe Thr
 385 390 395 400
 Thr Arg Ser Phe Asp Ser Glu Phe Lys Leu Ala Cys Lys Val Asp Gly
 405 410 415
 15 His Lys Asp Ile Gln Met Pro Asp Gly Met Gln Ala Arg Gly Val Cys
 420 425 430
 Ala Val Gly Gly Val Ala Pro Arg His Pro Asp Ala Leu Leu Ala Gly
 435 440 445
 20 Pro Ala Gln Gln Arg Arg Glu Ser Pro Pro Tyr His Thr Gly Cys Gly
 450 455 460
 His Asp Gln
 25 465
 (2) INFORMATION FOR SEQ ID NO: 366:
 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 152 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 366:
 Met Ala Asp Glu Ala Thr Arg Arg Val Val Ser Glu Ile Pro Val Leu
 1 5 10 15
 40 Lys Thr Asn Ala Gly Pro Arg Asp Arg Glu Leu Trp Val Gln Arg Leu
 20 25 30
 Lys Glu Glu Tyr Gln Ser Leu Ile Arg Tyr Val Glu Asn Asn Lys Asn
 35 40 45
 45 Ala Asp Asn Asp Trp Phe Arg Leu Glu Ser Asn Lys Glu Gly Thr Arg
 50 55 60
 Trp Phe Gly Lys Cys Trp Tyr Ile His Asp Leu Leu Lys Tyr Glu Phe
 50 65 70 75 80
 Asp Ile Glu Phe Asp Ile Pro Ile Thr Tyr Pro Thr Thr Ala Pro Glu
 85 90 95
 55 Ile Ala Val Pro Glu Leu Asp Gly Lys Thr Ala Lys Met Tyr Arg Gly
 100 105 110
 Gly Lys Ile Cys Leu Thr Asp His Phe Lys Pro Leu Trp Gly Gln Glu
 115 120 125
 60

Cys Ala Gln Ile Trp Thr Ser Ser Ser His Gly Ser Gly Ala Gly Ser
130 135 140

5 Met Xaa Gly Ser Gly Asn Pro Xaa
145 150

(2) INFORMATION FOR SEQ ID NO: 367:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 373 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 367:

Met Tyr Asp Gly Thr Lys Glu Val Pro Met Asn Pro Val Lys Ile Tyr
1 5 10 15

20 Gln Val Cys Asp Ile Pro Gln Pro Gln Gly Ser Ile Ile Asn Pro Gly
20 25 30

Ser Thr Gly Ser Ala Pro Trp Asp Glu Lys Asp Asn Asp Val Asp Glu
35 40 45

25

Glu Asp Glu Glu Asp Glu Leu Asp Gln Ser Gln His His Val Pro Ile
50 55 60

30 Gln Asp Thr Phe Pro Phe Leu Asn Ile Asn Gly Ser Pro Met Ala Pro
65 70 75 80

Ala Ser Val Gly Asn Cys Ser Val Gly Asn Cys Ser Pro Glu Ala Val
85 90 95

35 Trp Pro Lys Thr Glu Pro Leu Glu Met Glu Val Pro Gln Ala Pro Ile
100 105 110

Gln Pro Phe Tyr Ser Ser Pro Glu Leu Trp Ile Ser Ser Leu Pro Met
115 120 125

40

Thr Asp Leu Asp Ile Lys Phe Gln Tyr Arg Gly Lys Glu Tyr Gly Gln
130 135 140

45 Thr Met Thr Val Ser Asn Pro Gln Gly Cys Arg Leu Phe Tyr Gly Asp
145 150 155 160

Leu Gly Pro Met Pro Asp Gln Glu Glu Leu Phe Gly Pro Val Xaa Leu
165 170 175

50 Glu Gln Val Lys Phe Pro Gly Pro Glu His Ile Thr Asn Glu Lys Gln
180 185 190

Lys Leu Phe Thr Ser Lys Leu Leu Asp Val Met Asp Arg Gly Leu Ile
195 200 205

55

Leu Glu Val Ser Gly His Ala Ile Tyr Ala Ile Arg Leu Cys Gln Cys
210 215 220

60 Lys Val Tyr Trp Ser Gly Pro Cys Ala Pro Ser Leu Val Ala Pro Asn
225 230 235 240

Leu Ile Glu Arg Gln Lys Lys Val Lys Leu Phe Cys Leu Glu Thr Phe
 245 250 255
 5 Leu Ser Asp Leu Ile Ala His Gln Lys Gly Gln Ile Glu Lys Gln Pro
 260 265 270
 Pro Phe Glu Ile Tyr Leu Cys Phe Gly Glu Glu Trp Pro Asp Gly Lys
 275 280 285
 10 Pro Leu Glu Arg Lys Leu Ile Leu Val Gln Val Ile Pro Val Val Ala
 290 295 300
 Arg Met Ile Tyr Glu Met Phe Ser Gly Asp Phe Thr Arg Ser Phe Asp
 15 305 310 315 320
 Ser Gly Ser Val Arg Leu Gln Ile Ser Thr Pro Asp Ile Lys Asp Asn
 325 330 335
 20 Ile Val Ala Gln Leu Lys Gln Leu Tyr Arg Ile Leu Gln Thr Gln Glu
 340 345 350
 Ser Trp Gln Pro Met Gln Pro Thr Pro Ser Met Gln Leu Pro Pro Ala
 355 360 365
 25 Leu Pro Pro Gln Xaa
 370

30

(2) INFORMATION FOR SEQ ID NO: 368:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 83 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 368:

40 Met Gly Ser Ser Val Leu Pro Phe Cys Val Cys Val Thr Ser Pro Ser
 1 5 10 15
 Leu Gly Gly Arg Cys Ile Gln Gly Arg Phe Ala Ser His Ser Lys Phe
 20 25 30
 45 Trp Gly Phe Gly Arg Lys Thr Ala Ser Phe Gly Ala Val Gly Glu Thr
 35 40 45
 Pro Pro Asp Gln Glu Pro Gln Lys Glu Thr Glu Pro Ala Thr Ser Ser
 50 55 60
 50 His Ala Arg Pro Trp Ala Arg Val Ile Gly Leu Arg Ile Trp Pro Gln
 65 70 75 80
 Pro Asn Xaa
 55

60

(2) INFORMATION FOR SEQ ID NO: 369:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 369:

Met Leu Leu Ser Val Ala Ile Phe Ile Leu Leu Thr Leu Val Tyr Ala
 1 5 10 15

10 Tyr Trp Thr Met Xaa
 20

15 (2) INFORMATION FOR SEQ ID NO: 370:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 227 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 370:

Met Gly Ala Ser Ala Arg Leu Leu Arg Ala Val Ile Met Gly Ala Pro
 1 5 10 15

25 Gly Ser Gly Lys Gly Thr Val Ser Ser Arg Ile Thr Thr His Phe Glu
 20 25 30

30 Leu Lys His Leu Ser Ser Gly Asp Leu Leu Arg Asp Asn Met Leu Arg
 35 40 45

Gly Thr Glu Ile Gly Val Leu Ala Lys Ala Phe Ile Asp Gln Gly Lys
 50 55 60

35 Leu Ile Pro Asp Asp Val Met Thr Arg Leu Ala Leu His Glu Leu Lys
 65 70 75 80

Asn Leu Thr Gln Tyr Ser Trp Leu Leu Asp Gly Phe Pro Arg Thr Leu
 85 90 95

40 Pro Gln Ala Glu Ala Leu Asp Arg Ala Tyr Gln Ile Asp Thr Val Ile
 100 105 110

45 Asn Leu Asn Val Pro Phe Glu Val Ile Lys Gln Arg Leu Thr Ala Arg
 115 120 125

Trp Ile His Pro Ala Ser Gly Arg Val Tyr Asn Ile Glu Phe Asn Pro
 130 135 140

50 Pro Lys Thr Val Gly Ile Asp Asp Leu Thr Gly Glu Pro Leu Ile Gln
 145 150 155 160

Arg Glu Asp Asp Lys Pro Glu Thr Val Ile Lys Arg Leu Lys Ala Tyr
 165 170 175

55 Glu Asp Gln Thr Lys Pro Val Leu Glu Tyr Tyr Gln Lys Lys Gly Val
 180 185 190

60 Leu Glu Thr Phe Ser Gly Thr Glu Thr Asn Lys Ile Trp Pro Tyr Val
 195 200 205

Tyr Ala Phe Leu Gln Thr Lys Val Pro Gln Arg Ser Gln Lys Ala Ser
 210 215 220

5 Val Thr Pro
 225

10 (2) INFORMATION FOR SEQ ID NO: 371:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 79 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 371:

Met Phe Leu Asn Cys Glu Ile Leu Glu Tyr Cys Tyr Tyr Leu Thr Gln
 1 5 10 15

20 Leu Lys Ile Ser Met Gly Lys Tyr Leu Ser Ile Pro Thr Val Leu Leu
 20 25 30

25 Lys Ile Ile Arg Cys Ser Ile Thr Ala Val Ser Asp Ser Ser Thr Ser
 35 40 45

Trp Ala Ile Lys Ala Gln Leu Lys Ile Glu Asn Lys Asp Leu Asp Asn
 50 55 60

30 Lys Thr Ala Lys Gly Gly Gly Gln Glu Ala Leu Thr Cys Thr Xaa
 65 70 75

35 (2) INFORMATION FOR SEQ ID NO: 372:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 51 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 372:

Met Arg Ala Val Phe Pro Cys Cys Pro Phe Leu Thr Leu Met Leu Pro
 1 5 10 15

45 Leu Leu Glu Cys Leu Val Gly Met Ile Met Cys Tyr Leu Gly Ile Ser
 20 25 30

50 Phe Thr Asp Thr Arg Lys Thr Ala Gly Leu Lys Lys Lys Lys Lys Lys
 35 40 45

Lys Xaa Xaa
 50

55

(2) INFORMATION FOR SEQ ID NO: 373:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 61 amino acids

60

.577

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 373:

5 Met Phe Leu Met Arg Met His Leu Cys Phe Cys Lys Tyr Cys Cys Ser
1 5 10 15
Phe Ile Val Thr Pro Thr Ser Thr Ser Asn Thr Ala Ser Tyr Leu Trp
20 25 30
10 Pro Trp Ile Ser Ala Ser Met Ala Gly Arg Gly Ser Ser Trp Ala Cys
35 40 45
Thr Leu Asn Ala Val Thr Arg Glu Gly Leu Pro Glu Xaa
15 50 55 60

(2) INFORMATION FOR SEQ ID NO: 374:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 40 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 374:

Met Ser Leu Leu Asn Thr His Thr Leu Cys Phe Val Leu Phe Cys Phe
1 5 10 15
30 Thr Leu Ser Ile Asn Gln Glu Lys Leu Ala Asn His Leu Ala Phe Arg
20 25 30
Ile Leu Phe Phe Ile Val Phe Xaa
35 40
35

(2) INFORMATION FOR SEQ ID NO: 375:

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 44 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 375:

Met Cys Ser Gly Gln Ser Gln Val Trp Lys Met Ala Leu Gln Ala Leu
1 5 10 15
50 Asp Ser Glu Thr Val Val Ile Leu Pro Asp Met His Leu Ile Leu Ser
20 25 30
Leu Arg Leu Ile His Asn Ala Arg Pro Cys Leu Xaa
35 40
55

(2) INFORMATION FOR SEQ ID NO: 376:

60

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 203 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 376:

5 Met Leu Ile Ser Glu Glu Glu Ile Pro Phe Lys Asp Asp Pro Arg Asp
 1 5 10 15
 Glu Thr Tyr Lys Pro His Leu Glu Arg Glu Thr Pro Lys Pro Arg Arg
 20 25 30
 10 Lys Ser Gly Lys Val Lys Glu Glu Lys Glu Lys Lys Glu Ile Lys Val
 35 40 45
 Glu Val Glu Val Glu Val Lys Glu Glu Glu Asn Glu Ile Arg Glu Asp
 15 50 55 60
 Glu Glu Pro Pro Arg Lys Arg Gly Arg Arg Arg Lys Asp Asp Lys Ser
 65 70 75 80
 20 Pro Arg Leu Pro Lys Arg Arg Lys Lys Pro Pro Ile Gln Tyr Val Arg
 85 90 95
 Cys Glu Met Glu Gly Cys Gly Thr Val Leu Ala His Pro Arg Tyr Leu
 100 105 110
 25 Gln His His Ile Lys Tyr Gln His Leu Leu Lys Lys Lys Tyr Val Cys
 115 120 125
 Pro His Pro Ser Cys Gly Arg Leu Phe Arg Leu Gln Lys Gln Leu Leu
 130 135 140
 Arg His Ala Lys His His Thr Asp Gln Arg Asp Tyr Ile Cys Glu Tyr
 145 150 155 160
 35 Cys Ala Arg Ala Phe Lys Ser Ser His Asn Leu Ala Val His Arg Met
 165 170 175
 Ile His Thr Gly Glu Lys His Tyr Asn Val Arg Ser Val Asp Leu Leu
 180 185 190
 40 Val Asp Lys Arg His Leu Leu Ile Gly Thr Xaa
 195 200

45

(2) INFORMATION FOR SEQ ID NO: 377:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 377:

55 Met Leu Pro Arg Arg Thr Phe Tyr Phe Tyr Phe Ile Phe Ile Phe Phe
 1 5 10 15
 Leu Ala Ser Phe Trp Gly Phe Thr Leu Arg Ala Ser Phe
 20 25

60

(2) INFORMATION FOR SEQ ID NO: 378:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 136 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 378:

10 Met Phe Asp Ser Leu Ser Tyr Phe Lys Gly Ser Ser Leu Leu Leu Met
1 5 10 15
Leu Lys Thr Tyr Leu Ser Glu Asp Val Phe Gln His Ala Val Val Leu
20 25 30
15 Tyr Leu His Asn His Ser Tyr Ala Ser Ile Gln Ser Asp Asp Leu Trp
35 40 45
Asp Ser Phe Asn Glu Val Thr Asn Gln Thr Leu Asp Val Lys Arg Met
20 50 55 60
Met Lys Thr Trp Thr Leu Gln Lys Gly Phe Pro Leu Val Thr Val Gln
65 70 75 80
25 Lys Lys Gly Lys Glu Leu Phe Ile Gln Gln Glu Arg Phe Phe Leu Asn
85 90 95
Met Lys Pro Glu Ile Gln Pro Ser Asp Thr Arg Tyr Met Pro Ser Phe
100 105 110
30 Phe Ser Cys His Leu Phe Cys Thr Leu Arg Trp Lys Tyr Phe Glu Val
115 120 125
Phe Tyr Asn His Lys Phe Leu Xaa
35 130 135

(2) INFORMATION FOR SEQ ID NO: 379:

40

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 379:

Met Ala Trp Arg Arg Arg Glu Pro Ala Ser Gly Leu Ala Ala Cys Trp
1 5 10 15
50 Leu Trp Arg Cys Ser Pro Trp Pro Cys Ala Cys Pro Gly Pro Gly Ala
20 25 30
Gly Leu Ser Ser Gly Ser Arg Pro Trp
35 40
55

(2) INFORMATION FOR SEQ ID NO: 380:

60

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 468 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 380:

5 Met Glu Phe Leu Lys Val Ala Arg Arg Asn Lys Arg Glu Gln Leu Glu
 1 5 10 15
 10 Gln Ile Gln Lys Glu Leu Ser Val Leu Glu Glu Asp Ile Lys Arg Val
 20 25 30
 Glu Glu Met Ser Gly Leu Tyr Ser Pro Val Ser Glu Asp Ser Thr Val
 35 40 45
 15 Pro Gln Phe Glu Ala Pro Ser Pro Ser His Ser Ser Ile Ile Asp Ser
 50 55 60
 Thr Glu Tyr Ser Gln Pro Pro Gly Phe Ser Gly Ser Ser Gln Thr Lys
 65 70 75 80
 20 Lys Gln Pro Trp Tyr Asn Ser Thr Leu Ala Ser Arg Arg Lys Arg Leu
 85 90 95
 Thr Ala His Phe Glu Asp Leu Glu Gln Cys Tyr Phe Ser Thr Arg Met
 100 105 110
 Ser Arg Ile Ser Asp Asp Ser Arg Thr Ala Ser Gln Leu Asp Glu Phe
 115 120 125
 30 Gln Glu Cys Leu Ser Lys Phe Thr Arg Tyr Asn Ser Val Arg Pro Leu
 130 135 140
 Ala Thr Leu Ser Tyr Ala Ser Asp Leu Tyr Asn Gly Ser Ser Ile Val
 145 150 155 160
 35 Ser Ser Ile Glu Phe Asp Arg Asp Cys Asp Tyr Phe Ala Ile Ala Gly
 165 170 175
 Val Thr Lys Lys Ile Lys Val Tyr Glu Tyr Asp Thr Val Ile Gln Asp
 180 185 190
 Ala Val Asp Ile His Tyr Pro Glu Asn Glu Met Thr Cys Asn Ser Lys
 195 200 205
 45 Ile Ser Cys Ile Ser Trp Ser Ser Tyr His Lys Asn Leu Leu Ala Ser
 210 215 220
 Ser Asp Tyr Glu Gly Thr Val Ile Leu Trp Asp Gly Phe Thr Gly Gln
 225 230 235 240
 50 Arg Ser Lys Val Tyr Gln Glu His Glu Lys Arg Cys Trp Ser Val Asp
 245 250 255
 Phe Asn Leu Met Asp Pro Lys Leu Leu Ala Ser Gly Ser Asp Asp Ala
 260 265 270
 Lys Val Lys Leu Trp Ser Thr Asn Leu Asp Asn Ser Val Ala Ser Ile
 275 280 285
 60 Glu Ala Lys Ala Asn Val Cys Cys Val Lys Phe Ser Pro Ser Ser Arg

290 295 300

Tyr His Leu Ala Phe Gly Cys Ala Asp His Cys Val His Tyr Tyr Asp
305 310 315 320

5 Leu Arg Asn Thr Lys Gln Pro Ile Met Val Phe Lys Gly His Arg Lys
325 330 335

10 Ala Val Ser Tyr Ala Lys Phe Val Ser Gly Glu Glu Ile Val Ser Ala
340 345 350

Ser Thr Asp Ser Gln Leu Lys Leu Trp Asn Val Gly Lys Pro Tyr Cys
355 360 365

15 Leu Arg Ser Phe Lys Gly His Ile Asn Glu Lys Asn Phe Val Gly Leu
370 375 380

20 Ala Ser Asn Gly Asp Tyr Ile Ala Cys Gly Ser Glu Asn Asn Ser Leu
385 390 395 400

Tyr Leu Tyr Tyr Lys Gly Leu Ser Lys Thr Leu Leu Thr Phe Lys Phe
405 410 415

25 Asp Thr Val Lys Ser Val Leu Asp Lys Asp Arg Lys Glu Asp Asp Thr
420 425 430

Asn Glu Phe Val Ser Ala Val Cys Trp Arg Ala Leu Pro Asp Gly Glu
435 440 445

30 Ser Asn Val Leu Ile Ala Ala Asn Ser Gln Gly Thr Ile Lys Val Leu
450 455 460

Glu Leu Val Xaa
465

35

(2) INFORMATION FOR SEQ ID NO: 381:

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 381:

Met Arg Lys Glu Asp Gly Phe Trp Phe Phe Phe Phe Leu Phe Phe Phe
1 5 10 15

50

Val Val Gly Ser Lys Phe Val Asn Gly Asn Lys Leu Val
20 25

55

(2) INFORMATION FOR SEQ ID NO: 382:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 382:

Met Pro Leu Ala Pro Tyr Cys Asp Leu Leu Val Ala Leu Ser Phe Ala
1 5 10 15

5 Leu Val Leu Glu Ser Pro Val Asp Ser Ser Asp Phe Thr
20 25

10 (2) INFORMATION FOR SEQ ID NO: 383:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 138 amino acids

(B) TYPE: amino acid

15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 383:

Met Asn Ser Leu Val Ser Trp Gln Leu Leu Leu Phe Leu Cys Ala Thr
1 5 10 15

20 His Phe Gly Glu Pro Leu Glu Lys Val Ala Ser Val Gly Asn Ser Arg
20 25 30

25 Pro Thr Gly Gln Gln Leu Glu Ser Leu Gly Leu Leu Ala Pro Gly Glu
35 40 45

Gln Ser Leu Pro Cys Thr Glu Arg Lys Pro Ala Ala Thr Ala Arg Leu
50 55 60

30 Ser Arg Arg Gly Thr Ser Leu Ser Pro Pro Pro Glu Ser Ser Gly Ser
65 70 75 80

Pro Gln Gln Pro Gly Leu Ser Ala Pro His Ser Arg Gln Ile Pro Ala
85 90 95

35 Pro Gln Gly Ala Val Leu Val Gln Arg Glu Lys Asp Leu Pro Asn Tyr
100 105 110

40 Asn Trp Asn Ser Phe Gly Leu Arg Phe Gly Lys Arg Glu Ala Ala Pro
115 120 125

Gly Asn His Gly Arg Ser Ala Gly Arg Gly
130 135

45

(2) INFORMATION FOR SEQ ID NO: 384:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 74 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 384:

55 Met Ser Cys Phe Ile Asp Ser Xaa Asp Ser Lys Ile Leu His Leu Leu
1 5 10 15

Val Val Ser Phe Ile Cys Xaa Leu Phe Leu Leu Ile Leu Thr His Gly
20 25 30

60

Ile Leu Ile Leu Arg Xaa Phe Phe Ser Val Xaa Xaa His Ser Leu Lys
 35 40 45
 5 Asn Asn Leu Glu Glu Tyr Leu Ile Leu Met Asn Lys Ala Leu Leu Thr
 50 55 60
 Arg Glu Asp Phe Phe Val Leu Pro Xaa Ala
 65 70
 10
 (2) INFORMATION FOR SEQ ID NO: 385:
 (i) SEQUENCE CHARACTERISTICS:
 15 (A) LENGTH: 521 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 385:
 20 Met Ser Ala Gly Glu Val Glu Arg Leu Val Ser Glu Leu Ser Gly Gly
 1 5 10 15
 Thr Gly Gly Asp Glu Glu Glu Glu Trp Leu Tyr Gly Asp Glu Asn Glu
 20 25 30
 25 Val Glu Arg Pro Glu Glu Glu Asn Ala Ser Ala Asn Pro Pro Ser Gly
 35 40 45
 30 Ile Glu Asp Glu Thr Ala Glu Asn Gly Val Pro Lys Pro Lys Val Thr
 50 55 60
 Glu Thr Glu Asp Asp Ser Asp Ser Asp Ser Asp Asp Asp Glu Asp Asp
 65 70 75 80
 35 Val His Val Thr Ile Gly Asp Ile Lys Thr Gly Ala Pro Gln Tyr Gly
 85 90 95
 Ser Tyr Gly Thr Ala Pro Val Asn Leu Asn Ile Lys Thr Gly Gly Arg
 100 105 110
 40 Val Tyr Gly Thr Thr Gly Thr Lys Val Lys Gly Val Asp Leu Asp Ala
 115 120 125
 Pro Gly Ser Ile Asn Gly Val Pro Leu Leu Glu Val Asp Leu Asp Ser
 45 130 135 140
 Phe Glu Asp Lys Pro Trp Arg Lys Pro Gly Ala Asp Leu Ser Asp Tyr
 145 150 155 160
 50 Phe Asn Tyr Gly Phe Asn Glu Asp Thr Trp Lys Ala Tyr Cys Glu Lys
 165 170 175
 Gln Lys Arg Ile Arg Met Gly Leu Glu Val Ile Pro Val Thr Ser Thr
 180 185 190
 55 Thr Asn Lys Ile Thr Val Gln Gln Gly Arg Thr Gly Asn Ser Glu Lys
 195 200 205
 60 Glu Thr Ala Leu Pro Ser Thr Lys Ala Glu Phe Thr Ser Pro Pro Ser
 210 215 220

Leu Phe Lys Thr Gly Leu Pro Pro Ser Arg Arg Leu Pro Gly Ala Ile
225 230 235 240

5 Asp Val Ile Gly Gln Thr Ile Thr Ile Ser Arg Val Glu Gly Arg Arg
245 250 255

Arg Ala Asn Glu Asn Ser Asn Ile Gln Val Leu Ser Glu Arg Ser Ala
260 265 270

10 Thr Glu Val Asp Asn Asn Phe Ser Lys Pro Pro Pro Phe Phe Pro Pro
275 280 285

15 Gly Ala Pro Pro Thr His Leu Pro Pro Pro Pro Phe Leu Pro Pro Pro
290 295 300

Pro Thr Val Ser Thr Ala Pro Pro Leu Ile Pro Pro Pro Gly Phe Pro
305 310 315 320

20 Pro Pro Pro Gly Ala Pro Pro Pro Ser Leu Ile Pro Thr Ile Glu Ser
325 330 335

Gly His Ser Ser Gly Tyr Asp Ser Arg Ser Ala Arg Ala Phe Pro Tyr
340 345 350

25 Gly Asn Val Ala Phe Pro His Leu Pro Gly Ser Ala Pro Ser Trp Pro
355 360 365

30 Ser Leu Val Asp Thr Ser Lys Gln Trp Asp Tyr Tyr Ala Arg Arg Glu
370 375 380

Lys Asp Arg Asp Arg Glu Arg Asp Arg Asp Arg Glu Arg Asp Arg Asp
385 390 395 400

35 Arg Asp Arg Glu Arg Glu Arg Thr Arg Glu Arg Glu Arg Glu Arg Asp
405 410 415

His Ser Pro Thr Pro Ser Val Phe Asn Ser Asp Glu Glu Arg Tyr Arg
420 425 430

40 Tyr Arg Glu Tyr Ala Glu Arg Gly Tyr Glu Arg His Arg Ala Ser Arg
435 440 445

Glu Lys Glu Glu Arg His Arg Glu Arg Arg His Arg Glu Lys Glu Glu
450 455 460

45 Thr Arg His Lys Ser Ser Arg Ser Asn Ser Arg Arg Arg His Glu Ser
465 470 475 480

50 Glu Glu Gly Asp Ser His Arg Arg His Lys His Lys Lys Ser Lys Arg
485 490 495

Ser Lys Glu Gly Lys Glu Ala Gly Ser Glu Pro Ala Pro Glu Gln Glu
500 505 510

55 Ser Thr Glu Ala Thr Pro Ala Glu Xaa
515 520

60

(2) INFORMATION FOR SEQ ID NO: 386:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 137 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 386:

5
 10 Met Asn Ser Arg Gly Ile Trp Leu Ala Tyr Ile Ile Leu Val Gly Leu
 1 5 10 15
 Leu His Met Val Leu Leu Ser Ile Pro Phe Phe Ser Ile Pro Val Val
 20 25 30
 15 Trp Thr Leu Thr Asn Val Ile His Asn Leu Ala Thr Tyr Val Phe Leu
 35 40 45
 His Thr Val Lys Gly Thr Pro Phe Glu Thr Pro Asp Gln Gly Lys Ala
 50 55 60
 20 Arg Leu Leu Thr His Trp Glu Gln Met Asp Tyr Gly Leu Gln Phe Thr
 65 70 75 80
 Ser Ser Arg Lys Phe Leu Ser Ile Ser Pro Ile Val Leu Tyr Leu Leu
 85 90 95
 Ala Ser Phe Tyr Thr Lys Tyr Asp Ala Ala His Phe Leu Ile Asn Thr
 100 105 110
 30 Ala Ser Leu Leu Ser Val Leu Leu Pro Lys Leu Pro Gln Phe His Gly
 115 120 125
 Val Arg Val Phe Gly Ile Asn Lys Tyr
 130 135
 35

(2) INFORMATION FOR SEQ ID NO: 387:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 186 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 387:

40
 45 Met Ala Ala Gln Lys Asp Gln Gln Lys Asp Ala Glu Ala Glu Gly Leu
 1 5 10 15
 Ser Gly Thr Thr Leu Leu Pro Lys Leu Ile Pro Ser Gly Ala Gly Arg
 20 25 30
 Glu Trp Leu Glu Arg Arg Arg Ala Thr Ile Arg Pro Trp Ser Thr Phe
 35 40 45
 55 Val Asp Gln Gln Arg Phe Ser Arg Pro Arg Asn Leu Gly Glu Leu Cys
 50 55 60
 Gln Arg Leu Val Arg Asn Val Glu Tyr Tyr Gln Ser Asn Tyr Val Phe
 65 70 75 80
 60

Val Phe Leu Gly Leu Ile Leu Tyr Cys Val Val Thr Ser Pro Met Leu
85 90 95

5 Leu Val Ala Leu Ala Val Phe Phe Gly Ala Cys Tyr Ile Leu Tyr Leu
100 105 110

Arg Thr Leu Glu Ser Lys Leu Val Leu Phe Gly Arg Glu Val Ser Pro
115 120 125

10 Ala His Gln Tyr Ala Leu Ala Gly Gly Ile Ser Phe Pro Phe Phe Trp
130 135 140

Leu Ala Gly Ala Gly Ser Ala Val Phe Trp Val Leu Gly Ala Thr Leu
145 150 155 160

15 Val Val Ile Gly Ser His Ala Ala Phe His Gln Ile Glu Ala Val Asp
165 170 175

Gly Glu Glu Leu Gln Met Glu Pro Val Xaa
180 185

20

(2) INFORMATION FOR SEQ ID NO: 388:

25

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 388:

Met
1

35

(2) INFORMATION FOR SEQ ID NO: 389:

40

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 299 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 389:

45

Met Leu Ser Ile Phe Tyr Phe Ala Ile Pro Val Gly Ser Gly Leu Gly
1 5 10 15

Tyr Ile Ala Gly Ser Lys Val Lys Asp Met Ala Gly Asp Trp His Trp
20 25 30

50

Ala Leu Arg Val Thr Pro Gly Leu Gly Val Val Ala Val Leu Leu Leu
35 40 45

55

Phe Leu Val Val Arg Glu Pro Pro Arg Gly Ala Val Glu Arg His Ser
50 55 60

Asp Leu Pro Pro Leu Asn Pro Thr Ser Trp Trp Ala Asp Leu Arg Ala
65 70 75 80

60

Leu Ala Arg Asn Pro Ser Phe Val Leu Ser Ser Leu Gly Phe Thr Ala

587

85 90 95

Val Ala Phe Val Thr Gly Ser Leu Ala Leu Trp Ala Pro Ala Phe Leu
100 105 110

5 Leu Arg Ser Arg Val Val Leu Gly Glu Thr Pro Pro Cys Leu Pro Gly
115 120 125

10 Asp Ser Cys Ser Ser Ser Asp Ser Leu Ile Phe Gly Leu Ile Thr Cys
130 135 140

Leu Thr Gly Val Leu Gly Val Gly Leu Gly Val Glu Ile Ser Arg Arg
145 150 155 160

15 Leu Arg His Ser Asn Pro Arg Ala Asp Pro Leu Val Cys Ala Thr Gly
165 170 175

Leu Leu Gly Ser Ala Pro Phe Leu Phe Leu Ser Leu Ala Cys Ala Arg
180 185 190

20 Gly Ser Ile Val Ala Thr Tyr Ile Phe Ile Phe Ile Gly Glu Thr Leu
195 200 205

Leu Ser Met Asn Trp Ala Ile Val Ala Asp Ile Leu Leu Tyr Val Val
210 215 220

Ile Pro Thr Arg Arg Ser Thr Ala Glu Ala Phe Gln Ile Val Leu Ser
225 230 235 240

30 His Leu Leu Gly Asp Ala Gly Ser Pro Tyr Leu Ile Gly Leu Ile Ser
245 250 255

Asp Arg Leu Arg Arg Asn Trp Pro Pro Ser Phe Leu Ser Glu Phe Arg
260 265 270

35 Ala Leu Gln Phe Ser Leu Met Leu Cys Ala Phe Val Gly Ala Leu Gly
275 280 285

Gly Ala Leu Pro Gly His Arg His Leu His Xaa
290 295

40

(2) INFORMATION FOR SEQ ID NO: 390:

45

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 49 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 390:

Met Gly Pro Gln Gly Trp Val Arg Pro Leu Lys Thr Ala Pro Lys Leu
1 5 10 15

55 Gly Glu Ala Ile Arg Leu Ile Leu Phe Leu Asn Phe Val Lys Gln Cys
20 25 30

Ile Ala Ser Val Asn Leu Cys Ile Leu Arg Leu Asn Ile Thr Pro Leu
35 40 45

60

Leu

5

(2) INFORMATION FOR SEQ ID NO: 391:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 61 amino acids

10

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 391:

15 Met Tyr Val Asn Tyr Gly Thr Arg Asn Tyr Ser Thr Glu Gly Pro Ala
1 5 10 15Ala Leu Leu Asp Gln Ala Lys Leu Ser Leu Leu Val Trp Val Leu Cys
20 25 3020 Phe Val Leu Leu Phe Val Cys Phe Cys Gly Leu Ser Tyr Val Val Ile
35 40 45Ala Gln Val Pro Val Gly Leu Leu Cys Ile Thr Glu Xaa
50 55 60

25

(2) INFORMATION FOR SEQ ID NO: 392:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 79 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 392:

35

Met Leu Trp Phe Ala Asn Phe Phe Thr Tyr Leu Phe Leu Ser Gln Ser
1 5 10 1540 Val Ala Phe Val His Ile Ser His Ile Gly Val Arg Gln Val Asn Thr
20 25 30Asn Cys Tyr Phe Ser Arg Lys Ser Tyr Cys Tyr Gly Ile Leu Asn Pro
35 40 4545 Ile Asn Cys Ile Lys Gly Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys
50 55 60Lys Lys Lys Lys Ile Pro Ala Gly Arg Xaa Leu Phe Pro Phe Gly
65 70 75

50

(2) INFORMATION FOR SEQ ID NO: 393:

55

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 393:

60

Met Pro Gly Ala Phe Ser Glu Thr Val Ile Asn Asp Leu Leu Ser Leu
1 5 10 15
Phe Leu Val Leu Pro Ala Glu Leu Ser Tyr Ser Thr Leu Ser Gly Val
5 20 25 30
Tyr Arg Asn Ala
35

10

(2) INFORMATION FOR SEQ ID NO: 394:

(i) SEQUENCE CHARACTERISTICS:
15 (A) LENGTH: 180 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 394:

20 Met Ala Gln Ser Arg Asp Gly Gly Asn Pro Phe Ala Glu Pro Ser Glu
1 5 10 15
Leu Asp Asn Pro Phe Gln Asp Pro Ala Val Ile Gln His Arg Pro Ser
20 25 30
25 Arg Gln Tyr Ala Thr Leu Asp Val Tyr Asn Pro Phe Glu Thr Arg Glu
35 40 45
30 Pro Pro Pro Ala Tyr Glu Pro Pro Ala Pro Ala Pro Leu Pro Pro Pro
50 55 60
Ser Ala Pro Ser Leu Gln Pro Ser Arg Lys Leu Ser Pro Thr Glu Pro
65 70 75 80
35 Lys Asn Tyr Gly Ser Tyr Ser Thr Gln Ala Ser Ala Ala Ala Ala Thr
85 90 95
Ala Glu Leu Leu Lys Lys Gln Glu Glu Leu Asn Arg Lys Ala Glu Glu
100 105 110
40 Leu Asp Arg Arg Ser Glu Ser Cys Ser Met Leu Pro Trp Xaa Ala Gln
115 120 125
45 Leu Leu Asp Arg Thr Ile Gly Pro Leu Tyr Leu Leu Phe Val Gln Phe
130 135 140
Ser Pro Ala Phe Ser Arg Thr Ser Pro Trp Arg Ser Pro Lys Asn Phe
145 150 155 160
50 Arg Arg Leu Tyr Pro Pro Cys Thr Thr Ser Gly Cys Ala Ala Arg Trp
165 170 175
Xaa Phe Ser Xaa
180

55

(2) INFORMATION FOR SEQ ID NO: 395:

60 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 395:

5

Met Pro Thr Pro Cys Thr Ser Leu Pro Ser Cys Cys Gln His Arg Ser
 1 5 10 15

10

Ile Thr Met Thr Leu
 20

(2) INFORMATION FOR SEQ ID NO: 396:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 396:

Met Pro Leu Phe Ile Pro Leu Ile Phe Phe Leu Ser Leu Leu His Cys
 1 5 10 15

25

Gln Ser Lys His Pro Ile Gln Met Ser Leu Cys Met Cys Val Asn Ile
 20 25 30

Ser Leu Val Trp Ser Pro Val Arg Trp Ile Phe Gly Ser Lys Gly Leu
 35 40 45

30

Phe Ser Val His Leu Gln Ser Ser Gln Arg Pro Ser
 50 55 60

35

(2) INFORMATION FOR SEQ ID NO: 397:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 152 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 397:

Met Ala Gly Pro Arg Pro Xaa Trp Arg Asp Gln Leu Leu Phe Met Ser
 1 5 10 15

45

Ile Ile Val Leu Val Ile Val Val Ile Cys Leu Met Leu Tyr Ala Leu
 20 25 30

50

Leu Trp Glu Ala Gly Asn Leu Thr Asp Leu Pro Asn Leu Arg Ile Gly
 35 40 45

Phe Tyr Asn Phe Cys Leu Trp Asn Glu Asp Thr Ser Thr Leu Gln Cys
 50 55 60

55

His Gln Phe Pro Glu Leu Glu Ala Leu Gly Val Pro Arg Val Gly Leu
 65 70 75 80

60

Gly Leu Ala Arg Leu Gly Val Tyr Gly Ser Leu Val Leu Thr Leu Phe
 85 90 95

Ala Pro Gln Pro Leu Leu Leu Ala Gln Cys Asn Xaa Asp Glu Arg Ala
100 105 110

5 Trp Arg Leu Ala Val Gly Phe Leu Ala Val Ser Ser Val Leu Leu Ala
115 120 125

Gly Gly Leu Gly Leu Phe Leu Ser Tyr Val Trp Asn Gly Ser Xaa Ser
130 135 140

10 Pro Ser Arg Gly Leu Gly Phe Xaa
145 150

15 (2) INFORMATION FOR SEQ ID NO: 398:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 480 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 398:

25 Met Ser Asp Gly Phe Asp Arg Ala Pro Gly Ala Gly Arg Gly Arg Xaa
1 5 10 15

Arg Gly Leu Gly Arg Gly Gly Gly Gly Pro Xaa Gly Gly Gly Phe Pro
20 25 30

30 Xaa Gly Xaa Xaa Pro Ala Glu Arg Xaa Arg His Gln Pro Pro Gln Pro
35 40 45

Lys Ala Pro Gly Phe Leu Gln Pro Xaa Pro Leu Arg Gln Pro Arg Thr
50 55 60

35 Thr Pro Pro Pro Gly Ala Gln Cys Glu Val Pro Ala Ser Pro Gln Arg
65 70 75 80

40 Pro Ser Arg Pro Gly Ala Leu Pro Glu Gln Thr Arg Pro Leu Arg Ala
85 90 95

Pro Pro Ser Ser Gln Asp Lys Ile Pro Gln Gln Asn Ser Glu Ser Ala
100 105 110

45 Met Ala Lys Pro Gln Val Val Val Ala Pro Val Leu Met Ser Lys Leu
115 120 125

Ser Val Asn Ala Pro Glu Phe Tyr Pro Ser Gly Tyr Ser Ser Ser Tyr
130 135 140

50 Thr Glu Ser Tyr Glu Asp Gly Cys Glu Asp Tyr Pro Thr Leu Ser Glu
145 150 155 160

Tyr Val Gln Asp Phe Leu Asn His Leu Thr Glu Gln Pro Gly Ser Phe
165 170 175

Glu Thr Glu Ile Glu Gln Phe Ala Glu Thr Leu Asn Gly Cys Val Thr
180 185 190

60 Thr Asp Asp Ala Leu Gln Glu Leu Val Glu Leu Ile Tyr Gln Gln Ala

592

	195	200	205
	Thr Ser Ile Pro Asn Phe Ser Tyr Met Gly Ala Arg Leu Cys Asn Tyr		
	210	215	220
5	Leu Ser His His Leu Thr Ile Ser Pro Gln Ser Gly Asn Phe Arg Gln		
	225	230	235 240
10	Leu Leu Leu Gln Arg Cys Arg Thr Glu Tyr Glu Val Lys Asp Gln Ala		
	245	250	255
	Ala Lys Gly Asp Glu Val Thr Arg Lys Arg Phe His Ala Phe Val Leu		
	260	265	270
15	Phe Leu Gly Glu Leu Tyr Leu Asn Leu Glu Ile Lys Gly Thr Asn Gly		
	275	280	285
	Gln Val Thr Arg Ala Asp Ile Leu Gln Val Gly Leu Arg Glu Leu Leu		
	290	295	300
20	Asn Ala Leu Phe Ser Asn Pro Met Asp Asp Asn Leu Ile Cys Ala Val		
	305	310	315 320
25	Lys Leu Leu Lys Leu Thr Gly Ser Val Leu Glu Asp Ala Trp Lys Glu		
	325	330	335
	Lys Gly Lys Met Asp Met Glu Glu Ile Ile Gln Arg Ile Glu Asn Val		
	340	345	350
30	Val Leu Asp Ala Asn Cys Ser Arg Asp Val Lys Gln Met Leu Leu Lys		
	355	360	365
	Leu Val Glu Leu Arg Ser Ser Asn Trp Gly Arg Val His Ala Thr Ser		
	370	375	380
35	Thr Tyr Arg Glu Ala Thr Pro Glu Asn Asp Pro Asn Tyr Phe Met Asn		
	385	390	395 400
40	Glu Pro Thr Phe Tyr Thr Ser Asp Gly Val Pro Phe Thr Ala Ala Asp		
	405	410	415
	Pro Asp Tyr Gln Glu Lys Tyr Gln Glu Leu Leu Glu Arg Glu Asp Phe		
	420	425	430
45	Phe Pro Asp Tyr Glu Glu Asn Gly Thr Asp Leu Ser Gly Ala Gly Asp		
	435	440	445
	Pro Tyr Leu Asp Asp Ile Asp Asp Glu Met Asp Pro Glu Ile Glu Glu		
	450	455	460
50	Ala Tyr Glu Lys Phe Cys Leu Glu Ser Glu Arg Lys Arg Lys Gln Xaa		
	465	470	475 480

55

(2) INFORMATION FOR SEQ ID NO: 399:

60

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 423 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 399:

Met Glu Pro Lys Thr Ile Thr Asp Ala Leu Ala Ser Ser Ile Ile Lys
 1 5 10 15
 10 Ser Val Leu Pro Asn Phe Leu Pro Tyr Asn Val Met Leu Tyr Ser Asp
 20 25 30
 Ala Pro Val Ser Glu Leu Ser Leu Glu Leu Leu Leu Gln Val Val
 35 40 45
 15 Leu Pro Ala Leu Leu Glu Gln Gly His Thr Arg Gln Trp Leu Lys Gly
 50 55 60
 Leu Val Arg Ala Trp Thr Val Thr Ala Gly Tyr Leu Leu Asp Leu His
 65 70 75 80
 Ser Tyr Leu Leu Gly Asp Gln Glu Glu Asn Glu Asn Ser Ala Asn Gln
 85 90 95
 25 Gln Val Asn Asn Asn Gln His Ala Arg Asn Asn Asn Ala Ile Pro Val
 100 105 110
 Val Gly Glu Gly Leu His Ala Ala His Gln Ala Ile Leu Gln Gln Gly
 115 120 125
 30 Gly Pro Val Gly Phe Gln Xaa Tyr Arg Arg Pro Leu Asn Phe Pro Leu
 130 135 140
 Arg Ile Phe Leu Leu Ile Val Phe Met Cys Ile Thr Leu Leu Ile Ala
 145 150 155 160
 Ser Leu Ile Cys Leu Thr Leu Pro Val Phe Ala Gly Arg Trp Leu Met
 165 170 175
 40 Ser Phe Trp Thr Gly Thr Ala Lys Ile His Glu Leu Tyr Thr Ala Ala
 180 185 190
 Cys Gly Leu Tyr Val Cys Trp Leu Thr Ile Arg Ala Val Thr Val Met
 195 200 205
 45 Val Ala Trp Met Pro Gln Gly Arg Arg Val Ile Phe Gln Lys Val Lys
 210 215 220
 Glu Trp Ser Leu Met Ile Met Lys Thr Leu Ile Val Ala Val Leu Leu
 225 230 235 240
 Ala Gly Val Val Pro Leu Leu Leu Gly Leu Leu Phe Glu Leu Val Ile
 245 250 255
 55 Val Ala Pro Leu Arg Val Pro Leu Asp Gln Thr Pro Leu Phe Tyr Pro
 260 265 270
 Trp Gln Asp Trp Ala Leu Gly Val Leu His Ala Lys Ile Ile Ala Ala
 275 280 285
 60

Ile Thr Leu Met Gly Pro Gln Trp Trp Leu Lys Thr Val Ile Glu Gln
 290 295 300

5 Val Tyr Ala Asn Gly Ile Arg Asn Ile Asp Leu His Tyr Ile Val Arg
 305 310 315 320

Lys Leu Ala Ala Pro Val Ile Ser Val Leu Leu Ser Leu Cys Val
 325 330 335

10 Pro Tyr Val Ile Ala Ser Gly Val Val Pro Leu Leu Gly Val Thr Ala
 340 345 350

Glu Met Gln Asn Leu Val His Arg Arg Ile Tyr Pro Phe Leu Leu Met
 355 360 365

15 Val Val Val Leu Met Ala Ile Leu Ser Phe Gln Val Arg Gln Phe Lys
 370 375 380

20 Arg Leu Tyr Glu His Ile Lys Asn Asp Lys Tyr Leu Val Gly Gln Arg
 385 390 395 400

Leu Val Asn Tyr Glu Arg Lys Ser Gly Lys Gln Gly Ser Ser Pro Pro
 405 410 415

25 Pro Pro Gln Ser Ser Gln Glu
 420

30 (2) INFORMATION FOR SEQ ID NO: 400:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 78 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 400:

Met Leu Arg Leu Asp Ile Ile Asn Ser Leu Val Thr Thr Val Phe Met
 1 5 10 15

40 Leu Ile Val Ser Val Leu Ala Leu Ile Pro Glu Thr Thr Thr Leu Thr
 20 25 30

Val Gly Gly Gly Val Phe Ala Leu Val Thr Ala Val Cys Cys Leu Ala
 35 40 45

Asp Gly Ala Leu Ile Tyr Arg Lys Leu Leu Phe Asn Pro Ser Gly Pro
 50 55 60

50 Tyr Gln Lys Lys Pro Val His Glu Lys Lys Glu Val Leu Xaa
 65 70 75

55 (2) INFORMATION FOR SEQ ID NO: 401:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 74 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 401:

Met Leu Lys Gln Val Met Phe Val Phe Ser Gly Met Gly Pro Arg Ser
 1 5 10 15
 5 His Cys Trp Gly Leu Pro Leu His Val Ala Pro Leu Cys Arg Gly His
 20 25 30
 10 Gln Ala Asp Ser Ser His Leu Leu Pro Leu Lys His Gln Gly Ala Trp
 35 40 45
 Asn Arg Asn Leu Ala Asn Gln Arg His Phe Phe Cys Pro Ser Ile Phe
 50 55 60
 15 His Thr Cys Pro Thr Val Leu Phe Phe Xaa
 65 70

(2) INFORMATION FOR SEQ ID NO: 402:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 402:

Ala Arg Thr Ile Leu Val Leu Tyr Leu Ser Leu Gln Arg Leu Glu Asn
 1 5 10 15
 30 Leu Ala Tyr His
 20

35

(2) INFORMATION FOR SEQ ID NO: 403:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 87 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 403:

Met Pro Leu Pro Ser Val Pro Ile Leu Gly Ile Phe Ser Phe Leu Ile
 1 5 10 15
 45 Pro Ser Ser Gln Gly Val Ser Tyr Thr Lys Leu Pro Ile Ser Ser Pro
 20 25 30
 50 Gln Tyr Ser Pro Phe Val Asn Asp His Phe Ser Phe Leu Asn Pro Phe
 35 40 45
 Pro Val Gln Ile His Thr Gly Phe Ala Arg Val Gly Ser Tyr Met Gln
 50 55 60
 55 Met Pro Leu Val His Leu Cys Leu Leu Gln Thr Ser Leu Met Lys Asn
 65 70 75 80
 60 Ser Gly Val Gln Gln Gly Ser
 85

(2) INFORMATION FOR SEQ ID NO: 404:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 92 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 404:

Met Asn Ala Ala Met Val His Ile Asn Arg Ala Leu Lys Leu Ile Ile
1 5 10 15

Arg Leu Phe Leu Val Glu Asp Leu Val Asp Ser Leu Lys Leu Ala Val
20 25 30

Phe Met Trp Leu Met Thr Tyr Val Gly Ala Val Phe Asn Gly Ile Thr
35 40 45

20

Leu Leu Ile Leu Ala Glu Leu Leu Ile Phe Ser Val Pro Ile Val Tyr
50 55 60

Glu Lys Tyr Lys Thr Gln Ile Asp His Tyr Val Gly Ile Ala Arg Asp
65 70 75 80

25

Gln Thr Lys Ser Ile Val Glu Lys Ile Pro Ser Lys
85 90

30

(2) INFORMATION FOR SEQ ID NO: 405:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 405:

Met Ala Cys Ser Cys Leu Met Ile Gln Ser Phe Ser Thr Ser Ala Leu
1 5 10 15

Val Leu Phe Tyr Gly
20

45

(2) INFORMATION FOR SEQ ID NO: 406:

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 174 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 406:

55

Met Glu Glu Gly Gly Asn Leu Gly Gly Leu Ile Lys Met Val His Leu
1 5 10 15

60

Leu Val Leu Ser Gly Ala Trp Gly Met Gln Met Trp Val Thr Phe Val
20 25 30

Ser Gly Phe Pro Ala Phe Pro Lys Pro Ser Pro Thr Tyr Leu Arg Thr
 35 40 45
 5 Ser Ala Glu Gln Thr Leu Pro Leu Leu Leu Pro His Leu His Gly Leu
 50 55 60
 Cys Leu His Gln Pro Leu His Leu Gly Phe Thr Ala Cys Leu Gly Ser
 65 70 75 80
 10 Ala His Ile Leu Gly Gly Gln Pro Ala Leu Pro Ala Val Pro Glu Pro
 85 90 95
 Tyr Ala Gly His Cys Gln Arg Pro Leu Ala Gly Thr Pro His His Ser
 100 105 110
 15 Cys His Val Gly Pro Ala Asn Arg Gly Arg Arg Ser Glu Ala Trp Val
 115 120 125
 Gly Arg Tyr Gln Ala Ala Asn Arg Phe Pro Ile Leu Asn Ala Xaa Cys
 130 135 140
 Glu Arg Arg Thr Pro Ser Thr Val Leu Ser Ala Arg Ile Ser Ser Ala
 145 150 155 160
 25 Thr Met Gly Cys Pro Leu Phe Ala Ile Trp Ala Ala Ser Xaa
 165 170

30

(2) INFORMATION FOR SEQ ID NO: 407:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 64 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 407:

35 Met Ala Phe Ile Leu Leu Phe Tyr Cys Leu Met Thr Phe Leu Ser Leu
 40 1 5 10 15
 Glu Gln Asn Ser Ala Thr Val Glu Pro Ser Ser His Glu Ile Leu His
 20 25 30
 45 Leu Leu Gln Asn Cys Phe Glu Leu Leu Arg Thr Ser Thr Ser Gln Cys
 35 40 45
 Thr Glu Gly Ile Pro Cys Gln Arg Tyr Gln Asn Gly Leu His Ile Xaa
 50 55 60
 50

55

(2) INFORMATION FOR SEQ ID NO: 408:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 280 amino acids
 (B) TYPE: amino acid

60

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 408:

5 Met Glu Ala Val Val Asn Leu Tyr Gln Glu Val Met Lys His Ala Asp
 1 5 10 15
 Pro Arg Ile Gln Gly Tyr Pro Leu Met Gly Ser Pro Leu Leu Met Thr
 20 25 30
 10 Ser Ile Leu Leu Thr Tyr Val Tyr Phe Val Leu Ser Leu Gly Pro Arg
 35 40 45
 Ile Met Ala Asn Arg Lys Pro Phe Gln Leu Arg Gly Phe Met Ile Val
 50 55 60
 15 Tyr Asn Phe Ser Leu Val Ala Leu Ser Leu Tyr Ile Val Tyr Glu Phe
 65 70 75 80
 Leu Met Ser Gly Trp Leu Ser Thr Tyr Thr Trp Arg Cys Asp Pro Val
 85 90 95
 Asp Tyr Ser Asn Ser Pro Glu Ala Leu Arg Met Val Arg Val Ala Trp
 100 105 110
 25 Leu Phe Leu Phe Ser Lys Phe Ile Glu Leu Met Asp Thr Val Ile Phe
 115 120 125
 Ile Leu Arg Lys Lys Asp Gly Gln Val Thr Phe Leu His Val Phe His
 130 135 140
 30 His Ser Val Leu Pro Trp Ser Trp Trp Trp Gly Val Lys Ile Ala Pro
 145 150 155 160
 Gly Gly Met Gly Ser Phe His Ala Met Ile Asn Ser Ser Val His Val
 165 170 175
 35 Ile Met Tyr Leu Tyr Tyr Gly Leu Ser Ala Phe Gly Pro Val Ala Gln
 180 185 190
 40 Pro Tyr Leu Trp Trp Lys Lys His Met Thr Ala Ile Gln Leu Ile Gln
 195 200 205
 Phe Val Leu Val Ser Leu His Ile Ser Gln Tyr Tyr Phe Met Ser Ser
 210 215 220
 45 Cys Asn Tyr Gln Tyr Pro Val Ile Ile His Leu Ile Trp Met Tyr Gly
 225 230 235 240
 Thr Ile Phe Phe Met Leu Phe Ser Asn Phe Trp Tyr His Ser Tyr Thr
 245 250 255
 50 Lys Gly Lys Arg Leu Pro Arg Ala Leu Gln Gln Asn Gly Ala Pro Gly
 260 265 270
 55 Ile Ala Lys Val Lys Ala Asn Xaa
 275 280

60 (2) INFORMATION FOR SEQ ID NO: 409:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 284 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 409:

Met Xaa Leu Trp Pro Gln Thr Cys Ser Gly Lys Phe Asp Gly Thr Leu
 1 5 10 15
 Ala Phe Ser Ile His Xaa Leu Ala Val Ile Leu Gly Asp Gln Leu Thr
 20 25 30
 Ala Ala Asp Leu Val Pro Ile Phe Asn Gly Phe Leu Lys Asp Leu Asp
 35 40 45
 Glu Val Arg Ile Gly Val Leu Lys His Leu His Asp Phe Leu Lys Leu
 50 55 60
 Leu His Ile Asp Lys Arg Arg Glu Tyr Leu Tyr Gln Leu Gln Glu Phe
 65 70 75 80
 Leu Val Thr Asp Asn Ser Arg Asn Trp Arg Phe Arg Ala Glu Leu Ala
 85 90 95
 Glu Gln Leu Ile Leu Leu Leu Glu Leu Tyr Ser Pro Arg Asp Val Tyr
 100 105 110
 Asp Tyr Leu Arg Pro Ile Ala Leu Asn Leu Cys Ala Asp Lys Val Ser
 115 120 125
 Ser Val Arg Trp Ile Ser Tyr Lys Leu Val Ser Glu Met Val Lys Lys
 130 135 140
 Leu His Ala Ala Thr Pro Pro Thr Phe Gly Val Asp Leu Ile Asn Glu
 145 150 155 160
 Leu Val Glu Asn Phe Gly Arg Cys Pro Lys Trp Ser Gly Arg Gln Ala
 165 170 175
 Phe Val Phe Val Cys Gln Thr Val Ile Glu Asp Asp Cys Leu Pro Met
 180 185 190
 Asp Gln Phe Ala Val His Leu Met Pro His Leu Leu Thr Leu Ala Asn
 195 200 205
 Asp Arg Val Pro Asn Val Arg Val Leu Leu Ala Lys Thr Leu Arg Gln
 210 215 220
 Thr Leu Leu Glu Lys Asp Tyr Phe Leu Ala Ser Ala Ser Cys His Gln
 225 230 235 240
 Glu Ala Val Glu Gln Thr Ile Met Ala Leu Gln Met Asp Arg Asp Ser
 245 250 255
 Asp Val Lys Tyr Phe Ala Ser Ile His Pro Ala Ser Thr Lys Ile Ser
 260 265 270
 Glu Asp Ala Met Ser Thr Ala Ser Ser Thr Tyr Xaa
 275 280

(2) INFORMATION FOR SEQ ID NO: 410:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 187 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 410:

Met Leu Phe Leu Phe Phe Val Ile Ile Phe Leu Phe Val Phe Leu Ile
 1 5 10 15
 Leu Ile Ile Gln Phe Ser Lys Pro Leu Thr Asn Pro His Pro Pro Ala
 20 25 30
 Gly Xaa Ser Asp Arg Arg Arg Arg Tyr Ser Ser Tyr Arg Ser His Asp
 35 40 45
 His Tyr Gln Arg Gln Arg Val Leu Gln Lys Glu Arg Ala Ile Glu Glu
 50 55 60
 Arg Arg Val Val Phe Ile Gly Lys Ile Pro Gly Arg Met Thr Arg Ser
 65 70 75 80
 Glu Leu Lys Gln Arg Phe Ser Val Phe Gly Glu Ile Glu Glu Cys Thr
 85 90 95
 Ile His Phe Arg Val Gln Gly Asp Asn Tyr Gly Phe Val Thr Tyr Arg
 100 105 110
 Tyr Ala Glu Glu Ala Phe Ala Ala Ile Glu Ser Gly His Lys Leu Arg
 115 120 125
 Gln Ala Asp Glu Gln Pro Phe Asp Leu Cys Phe Gly Gly Arg Arg Xaa
 130 135 140
 Xaa Cys Lys Arg Ser Tyr Ser Asp Leu Asp Ser Asn Arg Glu Asp Phe
 145 150 155 160
 Asp Pro Ala Pro Val Lys Ser Lys Phe Asp Ser Leu Asp Phe Asp Thr
 165 170 175
 Leu Leu Lys Gln Ala Gln Lys Asn Leu Arg Arg
 180 185

50

(2) INFORMATION FOR SEQ ID NO: 411:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 237 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 411:

Met Lys Leu Pro Gly Lys Phe Arg Arg Ala His Gln Gly Asn Leu Glu
 1 5 10 15
 60

Ser Gln Leu Thr Ser Glu Ser Tyr Tyr Lys Glu Thr Leu Ser Val Pro
 20 25 30
 5 Thr Val Glu His Ile Ile Gln Glu Leu Lys Asp Ile Phe Ser Glu Gln
 35 40 45
 His Leu Lys Ala Leu Lys Cys Leu Ser Leu Val Pro Ser Val Met Gly
 50 55 60
 10 Gln Leu Lys Phe Asn Thr Ser Glu Glu His His Ala Asp Met Tyr Arg
 65 70 75 80
 Ser Asp Leu Pro Asn Pro Asp Thr Leu Ser Ala Glu Leu His Cys Trp
 85 90 95
 15 Arg Ile Lys Trp Lys His Arg Gly Lys Asp Ile Glu Leu Pro Ser Thr
 100 105 110
 Ile Tyr Glu Ala Leu His Leu Pro Asp Ile Lys Phe Phe Pro Asn Val
 115 120 125
 Tyr Ala Leu Leu Lys Val Leu Cys Ile Leu Pro Val Met Lys Val Glu
 130 135 140
 25 Asn Glu Arg Tyr Glu Asn Gly Arg Lys Arg Leu Lys Ala Tyr Leu Arg
 145 150 155 160
 Asn Thr Leu Thr Asp Gln Arg Ser Ser Asn Leu Ala Leu Leu Asn Ile
 165 170 175
 30 Asn Phe Asp Ile Lys His Asp Leu Asp Leu Met Val Asp Thr Tyr Ile
 180 185 190
 Lys Leu Tyr Thr Xaa Xaa Ser Xaa Leu Xaa Thr Xaa Xaa Ser Xaa Xaa
 195 200 205
 Val Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa
 210 215 220
 40 Asp Xaa Xaa Xaa Arg Glu Lys Ala Val Arg Cys Met Xaa
 225 230 235

45 (2) INFORMATION FOR SEQ ID NO: 412:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 192 amino acids

(B) TYPE: amino acid

50 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 412:

Met Lys Pro Met Ala Val Val Ala Ser Thr Val Leu Gly Leu Val Gln
 1 5 10 15
 55 Asn Met Arg Ala Phe Gly Gly Ile Leu Val Val Val Tyr Tyr Val Phe
 20 25 30
 Ala Ile Ile Gly Ile Asn Leu Phe Arg Gly Val Ile Val Ala Leu Pro
 35 40 45

Gly Asn Ser Ser Leu Ala Pro Ala Asn Gly Ser Ala Pro Cys Gly Ser
 50 55 60
 5 Phe Glu Gln Leu Glu Tyr Trp Ala Asn Asn Phe Asp Asp Phe Ala Ala
 65 70 75 80
 Ala Leu Val Thr Leu Trp Asn Leu Met Val Val Asn Asn Trp Gln Val
 85 90 95
 10 Phe Leu Asp Ala Tyr Arg Arg Tyr Ser Gly Pro Trp Ser Lys Ile Tyr
 100 105 110
 Phe Val Leu Trp Trp Leu Val Ser Ser Val Ile Trp Val Asn Leu Phe
 115 120 125
 15 Leu Ala Leu Ile Leu Glu Asn Phe Leu His Lys Trp Asp Pro Arg Ser
 130 135 140
 20 His Leu Gln Pro Leu Ala Gly Thr Pro Glu Ala Thr Tyr Gln Met Thr
 145 150 155 160
 Val Glu Leu Leu Phe Arg Asp Ile Leu Glu Glu Pro Gly Glu Asp Glu
 165 170 175
 25 Leu Thr Glu Arg Leu Ser Gln His Pro His Leu Trp Leu Cys Arg Xaa
 180 185 190

30

35 (2) INFORMATION FOR SEQ ID NO: 413:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 413:

Asn Val Val Val Val Ala Phe Gly Leu Ile Leu Ile Ile Glu Ser Leu
 1 5 10 15

45 Gly Glu Gln Cys Pro
 20

50 (2) INFORMATION FOR SEQ ID NO: 414:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 51 amino acids

(B) TYPE: amino acid

55 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 414:

Met Asn Trp Gly Leu Ser Ile Trp Leu His Tyr Tyr Glu Lys Lys Lys
 1 5 10 15

60

Glu Gln Val Phe Leu Val Ile Leu Ala His Val Val Arg Arg Cys Ala
 20 25 30

5 Ser Asp Gly Ile Leu Gln Phe Glu Ser Ser Leu Leu Lys Met Arg Arg
 35 40 45

Ala Pro Xaa
 50

10

(2) INFORMATION FOR SEQ ID NO: 415:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 415:

20 Met Leu Ile Ile Ser Leu Arg Pro Gln Phe Pro Ser Leu Ile Val Gln
 1 5 10 15

Leu Glu Cys Ser Val Leu Phe Leu Pro Ile Ser Leu Asn Leu Leu Leu
 20 25 30

25

30

(2) INFORMATION FOR SEQ ID NO: 416:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 163 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 416:

40 Met Val Lys Val Cys Asn Asp Ser Asp Arg Trp Ser Leu Ile Ser Leu
 1 5 10 15

Ser Asn Asn Ser Gly Lys Asn Val Glu Leu Lys Phe Val Asp Ser Leu
 20 25 30

45 Arg Arg Gln Phe Glu Phe Ser Val Asp Ser Phe Gln Ile Lys Leu Asp
 35 40 45

Ser Leu Leu Leu Phe Tyr Glu Cys Ser Glu Asn Pro Met Thr Glu Thr
 50 55 60

50

Phe His Pro Thr Ile Ile Gly Glu Ser Val Tyr Gly Asp Phe Gln Glu
 65 70 75 80

55 Ala Phe Asp His Leu Cys Asn Lys Ile Ile Ala Thr Arg Asn Pro Glu
 85 90 95

Glu Ile Arg Gly Gly Gly Leu Leu Lys Tyr Cys Asn Leu Leu Val Arg
 100 105 110

60 Gly Phe Arg Pro Ala Ser Asp Glu Ile Lys Thr Leu Gln Arg Tyr Met

115 120 125
Cys Ser Arg Phe Phe Ile Asp Phe Ser Asp Ile Gly Glu Gln Gln Arg
130 135 140
5
Lys Leu Glu Ser Tyr Leu Gln Asn His Phe Val Gly Ile Gly Arg Pro
145 150 155 160
10 Gln Val Xaa

15 (2) INFORMATION FOR SEQ ID NO: 417:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 174 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 417:

Met Ala Pro Lys Gly Lys Val Gly Thr Arg Gly Lys Lys Gln Ile Phe
1 5 10 15
25 Glu Glu Asn Arg Glu Thr Leu Lys Phe Tyr Leu Arg Ile Ile Leu Gly
20 25 30
Ala Asn Ala Ile Tyr Cys Leu Val Thr Leu Val Phe Phe Tyr Ser Ser
35 40 45
30 Ala Ser Phe Trp Ala Trp Leu Ala Leu Gly Phe Ser Leu Ala Val Tyr
50 55 60
35 Gly Ala Ser Tyr His Ser Met Ser Ser Met Ala Arg Ala Ala Phe Ser
65 70 75 80
Glu Asp Gly Ala Leu Met Asp Gly Gly Met Asp Leu Asn Met Glu Gln
85 90 95
40 Gly Met Ala Glu His Leu Lys Asp Val Ile Leu Leu Thr Ala Ile Val
100 105 110
Gln Val Leu Ser Cys Phe Ser Leu Tyr Val Trp Ser Phe Trp Leu Leu
115 120 125
45 Ala Pro Gly Arg Ala Leu Tyr Leu Leu Trp Val Asn Val Leu Gly Pro
130 135 140
50 Trp Phe Thr Ala Asp Ser Gly Thr Pro Ala Pro Glu His Asn Glu Lys
145 150 155 160
Arg Gln Arg Arg Gln Glu Arg Arg Gln Met Lys Arg Leu Xaa
165 170

55

(2) INFORMATION FOR SEQ ID NO: 418:

60 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 50 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 418:

5 Met Glu Leu Pro Lys Gly Leu Gln Gly Val Gly Pro Val Ala Met Met
 1 5 10 15

Arg Pro Phe Tyr Leu Leu Leu Pro Val Leu Cys Thr Gln Ala Leu Arg
 20 25 30

10 Gln Ser Gln Gly Lys Ser Pro Leu Leu Trp Lys Arg Thr Cys Cys Leu
 35 40 45

Ala Xaa
 15 50

(2) INFORMATION FOR SEQ ID NO: 419:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 120 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 419:

Met Leu Gly Lys Gly Gly Gly Arg Ala Gly Leu Leu Arg Tyr Arg Leu
 1 5 10 15

30 Leu Tyr Phe Thr Leu Val Val Gly Glu Gly Glu Pro Gly Glu Asn Lys
 20 25 30

Val Thr Ile Pro Phe Phe Glu Thr Gly Lys Lys Ile Ile Phe Cys Ser
 35 40 45

35 Val Lys Met Val Glu Asn Ser Asn Val Pro Ser His Lys Gly Pro Val
 50 55 60

40 Pro Leu Arg Ser Glu Gln Trp Glu Leu Lys Ile Ser Glu Thr Leu Gly
 65 70 75 80

Glu Gly Lys Ile Gly Phe Leu Leu Ile Gly Arg Cys Ser Ser Gly Xaa
 85 90 95

45 Gly Gly Leu Cys Phe Cys Trp Asp Val Leu Cys Cys Met Tyr Ala Tyr
 100 105 110

Met Asp Arg Ser Leu Leu Ser Leu
 115 120

50

(2) INFORMATION FOR SEQ ID NO: 420:

55

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 159 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 420:

Met Thr His Leu Leu Leu Thr Ala Thr Val Thr Pro Ser Glu Gln Asn
 1 5 10 15
 Ser Ser Arg Glu Pro Gly Trp Glu Thr Ala Met Ala Lys Asp Ile Leu
 5 20 25 30
 Gly Glu Ala Gly Leu His Phe Asp Glu Leu Asn Lys Leu Arg Val Leu
 35 40 45
 Asp Pro Glu Val Thr Gln Gln Thr Ile Glu Leu Lys Glu Glu Cys Lys
 10 50 55 60
 Asp Phe Val Asp Lys Ile Gly Gln Phe Gln Lys Ile Val Gly Gly Leu
 65 70 75 80
 Ile Glu Leu Val Asp Gln Leu Ala Lys Glu Ala Glu Asn Glu Lys Met
 15 85 90 95
 Lys Ala Ile Gly Ala Arg Asn Leu Leu Lys Ser Ile Ala Lys Gln Arg
 20 100 105 110
 Glu Ala Gln Gln Gln Gln Leu Gln Ala Leu Ile Ala Glu Lys Lys Met
 115 120 125
 Gln Leu Glu Arg Tyr Arg Val Glu Tyr Glu Ala Leu Cys Lys Val Glu
 25 130 135 140
 Ala Glu Gln Asn Glu Phe Ile Asp Gln Phe Ile Phe Gln Lys Xaa
 145 150 155
 30

(2) INFORMATION FOR SEQ ID NO: 421:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 154 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 421:
 Met Asn Val Gly Val Ala His Ser Glu Val Asn Pro Asn Thr Arg Val
 1 5 10 15
 Met Asn Ser Arg Gly Met Trp Leu Thr Tyr Ala Leu Gly Val Gly Leu
 45 20 25 30
 Leu His Ile Val Leu Leu Ser Ile Pro Phe Phe Ser Val Pro Val Ala
 35 40 45
 Trp Thr Leu Thr Asn Ile Ile His Asn Leu Gly Met Tyr Val Phe Leu
 50 55 60
 His Ala Val Lys Gly Thr Pro Phe Glu Thr Pro Asp Gln Gly Lys Ala
 65 70 75 80
 Arg Leu Leu Thr His Trp Glu Gln Leu Asp Tyr Gly Val Gln Phe Thr
 85 90 95
 Ser Ser Arg Lys Phe Phe Thr Ile Ser Pro Ile Ile Leu Tyr Phe Leu
 60 100 105 110

Ala Ser Phe Tyr Thr Lys Tyr Asp Pro Thr His Phe Ile Leu Asn Thr
115 120 125

5 Ala Ser Leu Leu Ser Val Leu Ile Pro Lys Met Pro Gln Leu His Gly
130 135 140

Val Arg Ile Phe Gly Ile Asn Lys Tyr Xaa
145 150

10

(2) INFORMATION FOR SEQ ID NO: 422:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 204 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 422:

20

Met Val Cys Gly Gly Phe Ala Cys Ser Lys Asn Cys Leu Cys Ala Leu
1 5 10 15

25

Asn Leu Leu Tyr Thr Leu Val Ser Leu Leu Ile Gly Ile Ala Ala
20 25 30

Trp Gly Ile Gly Phe Gly Leu Ile Ser Ser Leu Arg Val Val Gly Val
35 40 45

30

Val Ile Ala Val Gly Ile Phe Leu Phe Leu Ile Ala Leu Val Gly Leu
50 55 60

Ile Gly Ala Val Lys His His Gln Val Leu Leu Phe Phe Tyr Met Ile
65 70 75 80

35

Ile Leu Leu Leu Val Phe Ile Val Gln Phe Ser Val Ser Cys Ala Cys
85 90 95

40

Leu Ala Leu Asn Gln Glu Gln Gln Gly Gln Leu Leu Glu Val Gly Trp
100 105 110

Asn Asn Thr Ala Ser Ala Arg Asn Asp Ile Gln Arg Asn Leu Asn Cys
115 120 125

45

Cys Gly Phe Arg Ser Val Asn Pro Asn Asp Thr Cys Leu Ala Ser Cys
130 135 140

Val Lys Ser Asp His Ser Cys Ser Pro Cys Ala Pro Ile Ile Gly Glu
145 150 155 160

50

Tyr Ala Gly Glu Val Leu Arg Phe Val Gly Gly Ile Gly Leu Phe Phe
165 170 175

55

Ser Phe Thr Glu Ile Leu Gly Val Trp Leu Thr Tyr Arg Tyr Arg Asn
180 185 190

Gln Lys Asp Pro Arg Ala Asn Pro Ser Ala Phe Leu
195 200

60

(2) INFORMATION FOR SEQ ID NO: 423:

5 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 67 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 423:

10 Met Leu Gln Ser Ile Ile Lys Asn Ile Trp Ile Pro Met Lys Pro Tyr
1 5 10 15
Tyr Thr Lys Val Tyr Gln Glu Ile Trp Ile Gly Met Gly Leu Met Gly
20 25 30
15 Phe Ile Val Tyr Lys Ile Arg Ala Ala Asp Lys Arg Ser Lys Ala Leu
35 40 45
Lys Ala Ser Ala Pro Ala Pro Gly His His Asn Gln Ile Tyr Leu Glu
20 50 55 60
Tyr Met Xaa
65

25

(2) INFORMATION FOR SEQ ID NO: 424:

30 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 424:

35 Met Leu Gly Val Ser Leu Phe Leu Leu Val Val Leu Tyr His Tyr Val
1 5 10 15
Ala Val Asn Asn Pro Lys Lys Gln Glu
20 25
40

(2) INFORMATION FOR SEQ ID NO: 425:

45 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 299 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 425:

50

Met Ala Ala Xaa Glu Pro Ala Val Leu Ala Leu Pro Asn Ser Gly Ala
1 5 10 15
Gly Gly Ala Gly Ala Pro Ser Gly Thr Val Pro Val Leu Phe Cys Phe
20 25 30
55 Ser Val Phe Ala Arg Pro Ser Ser Val Pro His Gly Ala Gly Tyr Glu
35 40 45
60 Leu Leu Ile Gln Lys Phe Leu Ser Leu Tyr Gly Asp Gln Ile Asp Met

50 55 60
 His Arg Lys Phe Val Val Gln Leu Phe Ala Glu Glu Trp Gly Gln Tyr
 65 70 75 80
 5 Val Asp Leu Pro Lys Gly Phe Ala Val Ser Glu Arg Cys Lys Val Arg
 85 90 95
 10 Leu Val Pro Leu Gln Ile Gln Leu Thr Thr Leu Gly Asn Leu Thr Pro
 100 105 110
 Ser Ser Thr Val Phe Phe Cys Cys Asp Met Gln Glu Arg Phe Arg Pro
 115 120 125
 15 Ala Ile Lys Tyr Phe Gly Asp Ile Ile Ser Val Gly Gln Arg Leu Leu
 130 135 140
 Gln Gly Ala Arg Ile Leu Gly Ile Pro Val Ile Val Thr Glu Gln Tyr
 20 145 150 155 160
 Pro Lys Gly Leu Gly Ser Thr Val Gln Glu Ile Asp Leu Thr Gly Val
 165 170 175
 25 Lys Leu Val Leu Pro Lys Thr Lys Phe Ser Met Val Leu Pro Glu Val
 180 185 190
 Glu Ala Ala Leu Ala Glu Ile Pro Gly Val Arg Ser Val Val Leu Phe
 195 200 205
 30 Gly Val Glu Thr His Val Cys Ile Gln Gln Thr Ala Leu Glu Leu Val
 210 215 220
 Gly Arg Gly Val Glu Val His Ile Val Ala Asp Ala Thr Ser Ser Arg
 35 225 230 235 240
 Ser Met Met Asp Arg Met Phe Ala Leu Glu Arg Leu Ala Xaa Xaa Gly
 245 250 255
 40 Ile Ile Val Thr Thr Ser Glu Ala Val Leu Leu Gln Leu Val Ala Asp
 260 265 270
 Lys Asp His Pro Lys Phe Lys Glu Ile Gln Asn Leu Ile Lys Ala Ser
 275 280 285
 45 Ala Pro Glu Ser Gly Leu Leu Ser Lys Val Xaa
 290 295

50 (2) INFORMATION FOR SEQ ID NO: 426:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 amino acids

(B) TYPE: amino acid

55 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 426:

Met Arg Asp Leu Gly Thr Leu Leu Ser Pro Val Cys Ser
 1 5 10

60

(2) INFORMATION FOR SEQ ID NO: 427:

5 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 198 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 427:

Met Phe Gly Cys Leu Val Ala Gly Arg Leu Val Gln Thr Ala Ala Gln
1 5 10 15
Gln Val Ala Glu Asp Lys Phe Val Phe Asp Leu Pro Asp Tyr Glu Ser
15 20 25 30
Ile Asn His Val Val Val Phe Met Leu Gly Thr Ile Pro Phe Pro Glu
35 40 45
Gly Met Gly Gly Ser Val Tyr Phe Ser Tyr Pro Asp Ser Asn Gly Met
20 50 55 60
Pro Val Trp Gln Leu Leu Gly Phe Val Thr Asn Gly Lys Pro Ser Ala
25 65 70 75 80
Ile Phe Lys Ile Ser Gly Leu Lys Ser Gly Glu Gly Ser Gln His Pro
85 90 95
Phe Gly Ala Met Asn Ile Val Arg Thr Pro Ser Val Ala Gln Ile Gly
30 100 105 110
Ile Ser Val Glu Leu Leu Asp Ser Met Ala Gln Gln Thr Pro Val Gly
115 120 125
Asn Ala Ala Val Ser Ser Val Asp Ser Phe Thr Gln Phe Thr Gln Lys
35 130 135 140
Met Leu Asp Asn Phe Tyr Asn Phe Ala Ser Ser Phe Ala Val Ser Gln
40 145 150 155 160
Ala Gln Met Thr Pro Ser Pro Ser Glu Met Phe Ile Pro Ala Asn Val
165 170 175
Val Leu Lys Trp Tyr Glu Asn Phe Gln Arg Arg Leu Ala Gln Asn Pro
45 180 185 190
Xaa Phe Trp Xaa Thr Xaa
195

50

(2) INFORMATION FOR SEQ ID NO: 428:

55 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 428:

Met Gly Leu Pro Leu Met Ala Leu Met Trp Ser Thr Leu Pro Ala Ser

1 5 10 15

Ala Gly Val Asn Phe Ile Leu Ala Leu Pro Leu Leu Leu Leu Trp Lys
20 25 30

5

Asn Arg Gly Gly Val Gly Arg Ser Val Met Ser Ala Val Glu Xaa
35 40 45

10

(2) INFORMATION FOR SEQ ID NO: 429:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 370 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 429:

15

20 Met Lys Lys Val Glu Glu Lys Arg Val Asp Val Asn Ser Ala Val Ala
1 5 10 15

Met Gly Glu Val Ile Leu Ala Val Cys His Pro Asp Cys Ile Thr Thr
20 25 30

25 Ile Lys His Trp Ile Thr Ile Ile Arg Ala Arg Phe Glu Glu Val Leu
35 40 45

Thr Trp Ala Lys Gln His Gln Gln Arg Leu Glu Thr Ala Leu Ser Glu
50 55 60

30 Leu Val Ala Asn Ala Glu Leu Leu Glu Glu Leu Leu Ala Trp Ile Gln
65 70 75 80

35 Trp Ala Glu Thr Thr Leu Ile Gln Arg Asp Gln Glu Pro Ile Pro Gln
85 90 95

Asn Ile Asp Arg Val Lys Ala Leu Ile Ala Glu His Gln Thr Phe Met
100 105 110

40 Glu Glu Met Thr Arg Lys Gln Pro Asp Val Asp Arg Val Thr Lys Thr
115 120 125

Tyr Lys Arg Lys Asn Ile Glu Pro Thr His Ala Pro Phe Ile Glu Lys
130 135 140

45 Ser Arg Ser Gly Gly Arg Lys Ser Leu Ser Gln Pro Thr Pro Pro Pro
145 150 155 160

50 Met Pro Ile Leu Ser Gln Ser Glu Ala Lys Asn Pro Arg Ile Asn Gln
165 170 175

Leu Ser Ala Arg Trp Gln Gln Val Trp Leu Leu Ala Leu Glu Arg Gln
180 185 190

55 Arg Lys Leu Asn Asp Ala Leu Asp Arg Leu Glu Glu Leu Lys Glu Phe
195 200 205

Ala Asn Phe Asp Phe Asp Val Trp Arg Lys Lys Tyr Met Arg Trp Met
210 215 220

60

Asn His Lys Lys Ser Arg Val Met Asp Phe Phe Arg Arg Ile Asp Lys
 225 230 235 240
 5 Asp Gln Asp Gly Lys Ile Thr Arg Gln Glu Phe Ile Asp Gly Ile Leu
 245 250 255
 Ala Ser Lys Phe Pro Thr Thr Lys Leu Glu Met Thr Ala Val Ala Asp
 260 265 270
 10 Ile Phe Asp Arg Asp Gly Asp Gly Tyr Ile Asp Tyr Tyr Glu Phe Val
 275 280 285
 Ala Ala Leu His Pro Asn Lys Asp Ala Tyr Arg Pro Thr Thr Asp Ala
 290 295 300
 15 Asp Lys Ile Glu Asp Glu Val Thr Arg Gln Val Ala Gln Cys Lys Cys
 305 310 315 320
 Ala Lys Arg Phe Gln Val Glu Gln Ile Gly Glu Asn Lys Tyr Arg Phe
 20 325 330 335
 Phe Leu Gly Asn Gln Phe Gly Asp Ser Gln Gln Leu Arg Leu Val Arg
 340 345 350
 25 Ile Leu Arg Asn Arg Asp Gly Ser Arg Trp Trp Arg Met Asp Gly Leu
 355 360 365
 Gly Xaa
 30 370

(2) INFORMATION FOR SEQ ID NO: 430:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 430:
 Met Asn Val Lys Thr Phe Ser Xaa Asp His Met His Phe Leu Cys Cys
 1 5 10 15
 45 Leu Tyr Leu Arg Tyr Val Thr Phe Val Tyr Leu Asn Leu Phe
 20 25 30

(2) INFORMATION FOR SEQ ID NO: 431:

50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 431:
 Met Glu Pro His Leu Arg Cys Arg Val Thr Arg Val Arg Gly Ser Leu
 1 5 10 15
 60 Gly Asn Thr Gly Arg Trp Leu Leu

5 (2) INFORMATION FOR SEQ ID NO: 432:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 53 amino acids

(B) TYPE: amino acid

10 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 432:

Met His Tyr Leu Val Leu Gly Gly Leu Gly Val Phe Leu Phe Phe Ser
1 5 10 15
Cys Phe Val Phe Leu Phe Phe Xaa Phe Ser Phe Ala Phe Phe Pro Phe
20 25 30
Tyr Leu Glu Gly Met Gly Gly Ser Gly Asn Arg Glu Val Gly Gly Gly
35 40 45
Phe Cys Leu Phe Phe
50

25

(2) INFORMATION FOR SEQ ID NO: 433:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 176 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 433:

35 Met Val Ser Lys Ala Leu Leu Arg Leu Val Ser Ala Val Asn Arg Arg
1 5 10 15
Arg Met Lys Leu Leu Leu Gly Ile Ala Leu Leu Ala Tyr Val Ala Ser
20 25 30
40 Val Trp Gly Asn Phe Val Asn Met Arg Ser Ile Gln Glu Asn Gly Glu
35 40 45
Leu Lys Ile Glu Ser Lys Ile Glu Glu Met Val Glu Pro Leu Arg Glu
45 50 55 60
Lys Ile Arg Asp Leu Glu Lys Ser Phe Thr Gln Lys Tyr Pro Pro Val
65 70 75 80
50 Lys Phe Leu Ser Glu Lys Asp Arg Lys Arg Ile Leu Ile Thr Gly Gly
85 90 95
Ala Gly Phe Val Gly Ser His Leu Thr Asp Lys Leu Met Met Asp Gly
100 105 110
55 His Glu Val Thr Val Val Asp Asn Phe Phe Thr Gly Arg Lys Arg Asn
115 120 125
Val Glu His Trp Ile Gly His Glu Asn Phe Glu Leu Ile Asn His Asp
60 130 135 140

Val Trp Ser Pro Ser Thr Ser Arg Leu Thr Arg Tyr Thr Ile Trp His
 145 150 155 160

5 Leu Gln Pro Pro Leu Gln Thr Thr Cys Ile Ile Leu Ser Arg His Xaa
 165 170 175

10

(2) INFORMATION FOR SEQ ID NO: 434:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 77 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 434:

20

Met Leu Arg Cys Trp Pro Leu Phe Trp Leu Pro Leu Val Ser Pro Phe
 1 5 10 15

25

Cys Ser Leu Phe Trp Leu Leu Val Glu Trp Phe Gly Thr Asn Ile Asp
 20 25 30

Arg Glu Ser Tyr Asp Ala Ile Gly Gly Pro Ser Trp Met Thr Ala Ser
 35 40 45

30

Ser Phe Cys Leu Ser Asn Ser Asn Ile Trp Ser Leu Glu Ile Ser Ser
 50 55 60

Gly Ser Thr Ser Val Val His Ser Gln Gln Ala Met Asp
 65 70 75

35

(2) INFORMATION FOR SEQ ID NO: 435:

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 435:

45

Met Arg Ser Cys Glu Ile Gln Leu Cys Val Trp Leu Leu Val Ser Ser
 1 5 10 15

50

His Val Asp Met Val Leu Gly Gly Ser Pro Ser Thr Leu Tyr Met Met
 20 25 30

55

(2) INFORMATION FOR SEQ ID NO: 436:

60

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 436:

5 Met Val Val Asn Ser Leu Cys Phe Leu Ser Leu Leu Leu Val Ile Leu
1 5 10 15
Glu Leu Ser Thr Asp Ser Ser Ala Arg Leu Leu Tyr His Glu
20 25 30

10

(2) INFORMATION FOR SEQ ID NO: 437:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 69 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 437:

20

Met Asp Lys Gln Lys His Leu Glu Val Arg Arg Ser Val Phe Lys Ile
1 5 10 15

25

Gln Gly Lys Ile Ala Phe Ser Leu Met Phe Val Leu Lys Asp Leu Ser
20 25 30

Pro Thr Ile Phe Ser His Ser Ile Leu Leu Leu Leu Pro His His Val
35 40 45

30

Leu Pro Cys Thr Pro Gln Met Val Arg Gly Val Thr Gln Val Leu Arg
50 55 60

Glu Phe Gly Asp Gln
65

35

(2) INFORMATION FOR SEQ ID NO: 438:

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 438:

45

Met Pro Leu Cys Phe Phe Ser Phe Leu Cys Cys Trp Val Leu Val Phe
1 5 10 15

Lys Leu Ile

50

(2) INFORMATION FOR SEQ ID NO: 439:

55

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 439:

Met Lys Phe Ser Leu Val Leu Leu Ile Lys Ile Ile Ser Phe Glu Arg
 1 5 10 15
 5 Leu Leu Ile Phe Leu Phe Pro Leu Ser Phe Leu Pro Asn Ile Trp Arg
 20 25 30
 Arg Val Met Val Asn Leu Asn Ile Leu Phe Xaa
 35 40
 10

(2) INFORMATION FOR SEQ ID NO: 440:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 33 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 440:
 20 Met Leu Leu Phe Pro Ser Leu Leu Phe Ala Ala Thr Tyr Asn Val Ala
 1 5 10 15
 Asn Pro Ser Arg Leu Ile Leu Tyr Met Ile Ser Ala Gly Ala Asp Ser
 25 20 25 30
 Gln
 30

(2) INFORMATION FOR SEQ ID NO: 441:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 53 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 441:
 40 Met Trp Gln Val Arg Gly Leu Pro Pro Val Pro Leu Leu Leu Thr Met
 1 5 10 15
 Ser Pro Pro Pro Cys Leu Ser Ser Pro Phe Pro Phe Ile Ser Val Pro
 20 25 30
 45 Leu Phe Glu Ala Val Pro Ile Ser Val Ser Asp Gln Pro Ser Pro Xaa
 35 40 45
 Leu Thr Thr Leu Leu
 50 50

(2) INFORMATION FOR SEQ ID NO: 442:

55 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 64 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 442:

Met Ile Thr Ser Val Leu Val Phe Leu Ile Phe Phe Phe Pro Tyr Leu
 1 5 10 15
 5 Ser Leu Val Thr Leu Leu Gln Ala Arg Asn Leu Trp Val Ile His Arg
 20 25 30
 Ala Ala Leu Cys Glu Ser Gly Leu Phe His Trp Arg Lys Gly Ile Glu
 35 40 45
 10 Asn Gln Leu Glu Pro Met Tyr Phe Leu Pro His Gly Thr Leu Phe Leu
 50 55 60

15

- (2) INFORMATION FOR SEQ ID NO: 443:
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 34 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 443:

Met Leu Tyr Ser Cys Glu Pro Tyr Leu Ile Ile Leu Asn Ile Tyr Ser
 1 5 10 15
 30 Gln Lys Ala Phe Tyr Phe Tyr Phe Phe Glu Gly Ser Phe Ser Val Cys
 20 25 30
 Thr Leu

35

- (2) INFORMATION FOR SEQ ID NO: 444:
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 89 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 444:

Met Arg Gln Arg Gln Ala Ala Cys Gln Pro Pro Ser Arg Asn Gly
 1 5 10 15
 50 Leu Ala Gln Glu Cys Pro Pro His Ile Pro Ser Ser Phe Phe Leu Val
 20 25 30
 Lys Leu Leu Phe Ile Pro Trp Leu Ala Ser Leu Leu Ser Ser Pro Leu
 35 40 45
 55 Asn Leu Leu Leu Leu Val Ser Ile Ser Trp Asp Leu Gly Leu Lys Leu
 50 55 60
 Asn Leu Gln Gln Cys Arg Gln His Gln Val Leu Gln Glu Lys Asn Thr
 65 70 75 80

60

Lys Lys Phe Asn Lys Lys Lys Lys Lys
85

5

(2) INFORMATION FOR SEQ ID NO: 445:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 350 amino acids

10

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 445:

15	Met Asp Phe Ile Thr Ser Thr Ala Ile Leu Pro Leu Leu Phe Gly Cys	1	5	10	15
	Leu Gly Val Phe Gly Leu Phe Arg Leu Leu Gln Trp Val Arg Gly Lys	20	25	30	
20	Ala Tyr Leu Arg Asn Ala Val Val Val Ile Thr Gly Ala Thr Ser Gly	35	40	45	
	Leu Gly Lys Glu Cys Ala Lys Val Phe Tyr Ala Ala Gly Ala Lys Leu	50	55	60	
25	Val Leu Cys Gly Arg Asn Gly Gly Ala Leu Glu Glu Leu Ile Arg Glu	65	70	75	80
	Leu Thr Ala Ser His Ala Thr Lys Val Gln Thr His Lys Pro Tyr Leu	85	90	95	
30	Val Thr Phe Asp Leu Thr Asp Ser Gly Ala Ile Val Ala Ala Ala Ala	100	105	110	
35	Glu Ile Leu Gln Cys Phe Gly Tyr Val Asp Ile Leu Val Asn Asn Ala	115	120	125	
	Gly Ile Ser Tyr Arg Gly Thr Ile Met Asp Thr Thr Val Asp Val Asp	130	135	140	
40	Lys Arg Val Met Glu Thr Asn Tyr Phe Gly Pro Val Ala Leu Thr Lys	145	150	155	160
	Ala Leu Leu Pro Ser Met Ile Lys Arg Arg Gln Gly His Ile Val Ala	165	170	175	
45	Ile Ser Ser Ile Gln Gly Lys Met Ser Ile Pro Phe Arg Ser Ala Tyr	180	185	190	
50	Ala Ala Ser Lys His Ala Thr Gln Ala Phe Phe Asp Cys Leu Arg Ala	195	200	205	
	Glu Met Glu Gln Tyr Glu Ile Glu Val Thr Val Ile Ser Pro Gly Tyr	210	215	220	
55	Ile His Thr Asn Leu Ser Val Asn Ala Ile Thr Ala Asp Gly Ser Arg	225	230	235	240
60	Tyr Gly Val Met Asp Thr Thr Thr Ala Gln Gly Arg Ser Pro Val Glu	245	250	255	

Val Ala Gln Asp Val Leu Ala Ala Val Gly Lys Lys Lys Lys Asp Val
 260 265 270

5 Ile Leu Ala Asp Leu Leu Pro Ser Leu Ala Val Tyr Leu Arg Thr Leu
 275 280 285

Ala Pro Gly Leu Phe Phe Ser Leu Met Pro Pro Gly Pro Glu Lys Ser
 290 295 300

10 Gly Asn Pro Arg Thr Pro Ser Thr Leu Thr Ser Gln Gly Gln Gly Arg
 305 310 315 320

Glu Ala Ala Leu Leu Gly Leu Leu Thr Leu Gln Gly Thr Val Ala Phe
 15 325 330 335

Val Glu Thr Leu Met Glu Ile Cys Leu Thr Ser Gly Lys Asp
 340 345 350

20

(2) INFORMATION FOR SEQ ID NO: 446:

(i) SEQUENCE CHARACTERISTICS:
 25 (A) LENGTH: 49 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 446:

30 Met Val Phe Leu Pro Arg Gly Val Val Val Ser Gly Gly Ala Ala Cys
 1 5 10 15

Leu Trp Leu Thr Phe Ile Leu Glu Thr Glu Val Tyr Leu Asp Leu Ala
 20 25 30

35 Thr Glu Ala Arg Ala His Ser Arg Met Gly Leu Gly Leu Trp Pro Pro
 35 40 45

40 Asn

40

(2) INFORMATION FOR SEQ ID NO: 447:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 278 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 447:

Met Ala Ser Ala Glu Leu Asp Tyr Thr Ile Glu Ile Pro Asp Gln Pro
 1 5 10 15

55 Cys Trp Ser Gln Lys Asn Ser Pro Ser Pro Gly Gly Lys Glu Ala Glu
 20 25 30

Thr Arg Gln Pro Val Val Ile Leu Leu Gly Trp Gly Gly Cys Lys Asp
 35 40 45

60

Lys Asn Leu Ala Lys Tyr Ser Ala Ile Tyr His Lys Arg Gly Cys Ile
 50 55 60
 Val Ile Arg Tyr Thr Ala Pro Trp His Met Val Phe Phe Ser Glu Ser
 5 65 70 75 80
 Leu Gly Ile Pro Ser Leu Arg Val Leu Ala Gln Lys Leu Leu Glu Leu
 85 90 95
 10 Leu Phe Asp Tyr Glu Ile Glu Lys Glu Pro Leu Leu Phe His Val Phe
 100 105 110
 Ser Asn Gly Gly Val Met Leu Tyr Arg Tyr Val Leu Glu Leu Leu Gln
 115 120 125
 15 Thr Arg Arg Phe Cys Arg Leu Arg Val Val Gly Thr Ile Phe Asp Ser
 130 135 140
 Ala Pro Gly Asp Ser Asn Leu Val Gly Ala Leu Arg Ala Leu Ala Ala
 20 145 150 155 160
 Ile Leu Glu Arg Arg Ala Ala Met Leu Arg Leu Leu Leu Val Ala
 165 170 175
 25 Phe Ala Leu Val Val Val Leu Phe His Val Leu Leu Ala Pro Ile Thr
 180 185 190
 Ala Xaa Phe His Thr His Phe Tyr Asp Arg Leu Gln Asp Ala Gly Ser
 195 200 205
 30 Arg Trp Pro Glu Leu Tyr Leu Tyr Ser Arg Ala Asp Glu Val Val Leu
 210 215 220
 Ala Arg Asp Ile Glu Arg Met Val Glu Ala Arg Leu Ala Arg Arg Val
 35 225 230 235 240
 Leu Ala Arg Ser Val Asp Phe Val Ser Ser Ala His Val Ser His Leu
 245 250 255
 40 Arg Asp Tyr Pro Thr Tyr Tyr Thr Ser Leu Cys Val Asp Phe Met Arg
 260 265 270
 Asn Cys Val Arg Cys Xaa
 275
 45

(2) INFORMATION FOR SEQ ID NO: 448:

- 50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 199 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 448:

Met Ser Phe Ile Phe Asp Trp Ile Tyr Ser Gly Phe Ser Ser Val Leu
 1 5 10 15
 Gln Phe Leu Gly Leu Tyr Lys Lys Thr Gly Lys Leu Val Phe Leu Gly
 60 20 25 30

Leu Asp Asn Ala Gly Lys Thr Thr Leu Leu His Met Leu Lys Asp Asp
 35 40 45
 5 Arg Leu Gly Gln His Val Pro Thr Leu His Pro Thr Ser Glu Glu Leu
 50 55 60
 Thr Ile Ala Gly Met Thr Phe Thr Thr Phe Asp Leu Gly Gly His Val
 65 70 75 80
 10 Gln Ala Arg Arg Val Trp Lys Asn Tyr Leu Pro Ala Ile Asn Gly Ile
 85 90 95
 Val Phe Leu Val Asp Cys Ala Asp His Glu Arg Leu Leu Glu Ser Lys
 15 100 105 110
 Glu Glu Leu Asp Ser Leu Met Thr Asp Glu Thr Ile Ala Asn Val Pro
 115 120 125
 20 Ile Leu Ile Leu Gly Asn Lys Ile Asp Arg Pro Glu Ala Ile Ser Glu
 130 135 140
 Glu Arg Leu Arg Glu Met Phe Gly Leu Tyr Gly Gln Thr Thr Gly Lys
 145 150 155 160
 25 Gly Ser Ile Ser Leu Lys Glu Leu Asn Ala Arg Pro Leu Glu Val Phe
 165 170 175
 Met Cys Ser Val Leu Lys Arg Gln Gly Tyr Gly Glu Gly Phe Arg Trp
 30 180 185 190
 Met Ala Gln Tyr Ile Asp Xaa
 195

35

(2) INFORMATION FOR SEQ ID NO: 449:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 258 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 449:

45 Met Thr Leu Ser Arg Phe Ala Tyr Asn Gly Lys Arg Cys Pro Ser Ser
 1 5 10 15
 Tyr Asn Ile Leu Asp Asn Ser Lys Ile Ile Ser Glu Glu Cys Arg Lys
 20 25 30
 50 Glu Leu Thr Ala Leu Leu His His Tyr Tyr Pro Ile Glu Ile Asp Pro
 35 40 45
 His Arg Thr Val Lys Glu Lys Leu Pro His Met Val Glu Trp Trp Thr
 55 50 55 60
 Lys Ala His Asn Leu Leu Cys Gln Gln Lys Ile Gln Lys Phe Gln Ile
 65 70 75 80
 60 Ala Gln Val Val Arg Glu Ser Asn Ala Met Leu Arg Glu Gly Tyr Lys

	85	90	95
	Thr Phe Phe Asn Thr Leu Tyr His Asn Asn Ile Pro Leu Phe Ile Phe		
	100	105	110
5	Ser Ala Gly Ile Gly Asp Ile Leu Glu Glu Ile Ile Arg Gln Met Lys		
	115	120	125
	Val Phe His Pro Asn Ile His Ile Val Ser Asn Tyr Met Asp Phe Asn		
10	130	135	140
	Glu Asp Gly Phe Leu Gln Gly Phe Lys Gly Gln Leu Ile His Thr Tyr		
	145	150	155
15	Asn Lys Asn Ser Ser Val Cys Glu Asn Xaa Gly Tyr Phe Gln Gln Leu		
	165	170	175
	Glu Gly Lys Thr Asn Val Ile Leu Leu Gly Asp Ser Ile Gly Asp Leu		
20	180	185	190
	Thr Met Ala Asp Gly Val Pro Gly Val Gln Asn Ile Leu Lys Ile Gly		
	195	200	205
25	Phe Leu Asn Asp Lys Val Glu Glu Arg Arg Xaa Arg Tyr Met Asp Ser		
	210	215	220
	Tyr Asp Ile Val Leu Glu Lys Asp Glu Thr Leu Asp Val Val Asn Gly		
	225	230	235
30	Leu Leu Gln His Ile Leu Cys Gln Gly Val Gln Leu Glu Met Gln Gly		
	245	250	255
	Pro Xaa		

35

(2) INFORMATION FOR SEQ ID NO: 450:

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 87 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 450:

45

Met Ser His Val Leu Leu Cys Pro Ser Leu Ser Cys Ser Asn Leu Leu
1 5 10 15

50

Pro Pro Ser His Ser Leu Gly Thr Met Gly Ser Leu Ser Pro His Leu
20 25 30

Cys Gly His Thr Met Cys Pro Val Asn Pro Glu Leu Pro Leu Ser Ser
35 40 45

55

Arg Leu Thr Thr Asp Gln Pro Gln Pro Asp Ala Cys Ser Pro Thr Leu
50 55 60

60

Leu Thr Leu Pro Leu Pro Ser Ser Phe Leu Pro His Ser Lys Pro Thr
65 70 75 80

Phe Xaa His Pro Cys Ser Pro
85

5

(2) INFORMATION FOR SEQ ID NO: 451:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 315 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 451:

10

15

Met Phe Ser Ile Asn Pro Leu Glu Asn Leu Lys Val Tyr Ile Ser Ser
1 5 10 15

Arg Pro Pro Leu Val Val Phe Met Ile Ser Val Xaa Pro Met Ala Ile
20 25 30

20

Ala Phe Leu Thr Leu Gly Tyr Phe Phe Lys Ile Lys Glu Ile Lys Ser
35 40 45

25

Pro Glu Met Ala Glu Asp Trp Asn Thr Phe Leu Leu Arg Phe Asn Asp
50 55 60

Leu Asp Leu Cys Val Ser Glu Asn Glu Thr Leu Lys His Leu Thr Asn
65 70 75 80

30

Asp Thr Thr Thr Pro Glu Ser Thr Met Thr Ser Gly Gln Ala Arg Ala
85 90 95

Ser Thr Gln Ser Pro Gln Ala Leu Glu Asp Ser Gly Pro Val Asn Ile
100 105 110

35

Ser Val Ser Ile Thr Leu Thr Leu Asp Pro Leu Lys Pro Phe Gly Gly
115 120 125

Tyr Ser Arg Asn Val Thr His Leu Tyr Ser Thr Ile Leu Gly His Gln
130 135 140

40

Ile Gly Leu Ser Gly Arg Glu Ala His Glu Glu Ile Asn Ile Thr Phe
145 150 155 160

45

Thr Leu Pro Thr Ala Trp Ser Ser Asp Asp Cys Ala Leu His Gly His
165 170 175

Cys Glu Gln Val Val Phe Thr Ala Cys Met Thr Leu Thr Ala Ser Pro
180 185 190

50

Gly Val Phe Pro Val Thr Val Gln Pro Pro His Cys Val Pro Asp Thr
195 200 205

Tyr Ser Asn Ala Thr Leu Trp Tyr Lys Ile Phe Thr Thr Ala Arg Asp
210 215 220

55

Ala Asn Thr Lys Tyr Ala Gln Asp Tyr Asn Pro Phe Trp Cys Tyr Lys
225 230 235 240

60

Gly Ala Ile Gly Lys Val Tyr His Ala Leu Asn Pro Lys Leu Thr Val
245 250 255

Ile Val Pro Asp Asp Asp Arg Ser Leu Ile Asn Leu His Leu Met His
260 265 270

5 Thr Ser Tyr Phe Leu Phe Val Met Val Ile Thr Met Phe Cys Tyr Ala
275 280 285

Val Ile Lys Gly Arg Pro Ser Lys Leu Arg Gln Ser Asn Pro Glu Phe
290 295 300

10 Cys Pro Glu Lys Val Ala Leu Ala Glu Ala Xaa
305 310 315

15

(2) INFORMATION FOR SEQ ID NO: 452:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 52 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 452:

25 Met Pro Gly Leu Ser Leu Ala Leu Leu Pro Phe Gly Pro Gly Cys Thr
1 5 10 15

Glu Ala Leu His Ala Gly Cys Phe Pro Ala Phe Ala Ser Ala Thr Arg
20 25 30

30 Val Asn Gly Glu Ala Ala Leu Ser Pro Gly Leu Cys Asp Pro Ile Ser
35 40 45

Val Pro Tyr Val
50

35

(2) INFORMATION FOR SEQ ID NO: 453:

40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 383 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 453:

45 Met Ala Val Gly Gln Ile Met Thr Phe Gly Ser Pro Val Ile Gly Cys
1 5 10 15

50 Gly Phe Ile Ser Gly Trp Asn Leu Val Ser Met Cys Val Glu Tyr Val
20 25 30

Leu Leu Trp Lys Val Tyr Gln Lys Thr Pro Ala Leu Ala Val Lys Ala
35 40 45

55 Gly Leu Lys Glu Glu Glu Thr Glu Leu Lys Gln Leu Asn Leu His Lys
50 55 60

Asp Thr Glu Pro Lys Pro Leu Glu Gly Thr His Leu Met Gly Val Lys
65 70 75 80

60

Asp Ser Asn Ile His Glu Leu Glu His Glu Gln Glu Pro Thr Cys Ala
 85 90 95
 5 Ser Gln Met Ala Glu Pro Phe Arg Thr Phe Arg Asp Gly Trp Val Ser
 100 105 110
 Tyr Tyr Asn Gln Pro Val Phe Leu Ala Gly Met Gly Leu Ala Phe Leu
 115 120 125
 10 Tyr Met Thr Val Leu Gly Phe Asp Cys Ile Thr Thr Gly Tyr Ala Tyr
 130 135 140
 Thr Gln Gly Leu Ser Gly Phe His Pro Gln Tyr Phe Asp Gly Ser Ile
 145 150 155 160
 15 Ser Tyr Asn Trp Asn Asn Gly Asn Cys Ser Phe Tyr Leu Ala Thr Ser
 165 170 175
 20 Lys Met Trp Phe Gly Ser Ala Gly Leu Ile Ser Gly Leu Ala Gln Leu
 180 185 190
 Ser Cys Leu Ile Leu Cys Val Ile Ser Val Phe Met Pro Gly Ser Pro
 195 200 205
 25 Leu Asp Leu Ser Val Ser Pro Phe Glu Asp Ile Arg Ser Arg Phe Ile
 210 215 220
 Gln Gly Glu Ser Ile Thr Pro Thr Lys Ile Pro Glu Ile Thr Thr Glu
 225 230 235 240
 30 Ile Tyr Met Ser Asn Gly Ser Asn Ser Ala Asn Ile Val Pro Glu Thr
 245 250 255
 Ser Pro Glu Ser Val Pro Ile Ile Ser Val Ser Leu Leu Phe Ala Gly
 260 265 270
 Val Ile Ala Ala Arg Ile Gly Leu Trp Ser Phe Asp Leu Thr Val Thr
 275 280 285
 40 Gln Leu Leu Gln Glu Asn Val Ile Glu Ser Glu Arg Gly Ile Ile Asn
 290 295 300
 Gly Val Gln Asn Ser Met Asn Tyr Leu Leu Asp Leu Leu His Phe Ile
 305 310 315 320
 45 Met Val Ile Leu Ala Pro Asn Pro Glu Ala Phe Gly Leu Leu Val Leu
 325 330 335
 Ile Ser Val Ser Phe Val Ala Met Gly His Ile Met Tyr Phe Arg Phe
 340 345 350
 Ala Gln Asn Thr Leu Gly Asn Lys Leu Phe Ala Cys Gly Pro Asp Ala
 355 360 365
 55 Lys Glu Val Arg Lys Glu Asn Gln Ala Asn Thr Ser Val Val Xaa
 370 375 380

60 (2) INFORMATION FOR SEQ ID NO: 454:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 186 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 454:

5 Met Arg Ser Ile Gly Asn Lys Asn Thr Ile Leu Leu Gly Leu Gly Phe
 1 5 10 15
 10 Gln Ile Leu Gln Leu Ala Trp Tyr Gly Phe Gly Ser Glu Pro Trp Met
 20 25 30
 15 Met Trp Ala Ala Gly Ala Val Ala Ala Met Ser Ser Ile Thr Phe Pro
 35 40 45
 Ala Val Ser Ala Leu Val Ser Arg Thr Ala Asp Ala Asp Gln Gln Gly
 50 55 60
 20 Val Val Gln Gly Met Ile Thr Gly Ile Arg Gly Leu Cys Asn Gly Leu
 65 70 75 80
 Gly Pro Ala Leu Tyr Gly Phe Ile Phe Tyr Ile Phe His Val Glu Leu
 85 90 95
 25 Lys Glu Leu Pro Ile Thr Gly Thr Asp Leu Gly Thr Asn Thr Ser Pro
 100 105 110
 30 Gln His His Phe Glu Gln Asn Ser Ile Ile Pro Gly Pro Pro Phe Leu
 115 120 125
 Phe Gly Ala Cys Ser Val Leu Leu Ala Leu Leu Val Ala Leu Phe Ile
 130 135 140
 35 Pro Glu His Thr Asn Leu Ser Leu Arg Ser Ser Ser Trp Arg Lys His
 145 150 155 160
 Cys Gly Ser His Ser His Pro His Asn Thr Gln Ala Pro Gly Glu Ala
 165 170 175
 40 Lys Glu Pro Leu Leu Gln Asp Thr Asn Val
 180 185

(2) INFORMATION FOR SEQ ID NO: 455:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 163 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 455:

55 Met Leu Gln Thr Ser Asn Tyr Ser Leu Val Leu Ser Leu Gln Phe Leu
 1 5 10 15
 Leu Leu Ser Tyr Asp Leu Phe Val Asn Ser Phe Ser Glu Leu Leu Gln
 20 25 30
 60 Lys Thr Pro Val Ile Gln Leu Val Leu Phe Ile Ile Gln Asp Ile Ala

35 40 45
 Val Leu Phe Asn Ile Ile Ile Ile Phe Leu Met Phe Phe Asn Thr Phe
 50 55 60
 5 Val Phe Gln Ala Gly Leu Val Asn Leu Leu Phe His Lys Phe Lys Gly
 65 70 75 80
 10 Thr Ile Ile Leu Thr Ala Val Tyr Phe Ala Leu Ser Ile Ser Leu His
 85 90 95
 Val Trp Val Met Asn Leu Arg Trp Lys Asn Ser Asn Ser Phe Ile Trp
 100 105 110
 15 Thr Asp Gly Leu Gln Met Leu Phe Val Phe Gln Arg Leu Ala Ala Val
 115 120 125
 Leu Tyr Cys Tyr Phe Tyr Lys Arg Thr Ala Val Arg Leu Gly Asp Pro
 130 135 140
 20 His Phe Tyr Gln Asp Ser Leu Trp Leu Arg Lys Glu Phe Met Gln Val
 145 150 155 160
 Arg Arg Xaa
 25

30 (2) INFORMATION FOR SEQ ID NO: 456:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 46 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 456:

Met Arg Ile Gln Val Phe Ile Leu Leu Leu Gly Ala Gly Gly Thr Ser
 1 5 10 15
 40 Gln Phe Thr Lys Pro Pro Ser Leu Pro Leu Glu Pro Glu Pro Ala Val
 20 25 30
 Glu Ser Ser Pro Thr Glu Thr Ser Glu Gln Ile Arg Glu Lys
 35 40 45
 45

(2) INFORMATION FOR SEQ ID NO: 457:

50 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 105 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 457:

Met Ser Tyr Leu Ala Phe Leu Tyr Met Thr Phe Asp Phe Cys Cys Leu
 1 5 10 15
 60 Tyr Phe Ser Thr Val Tyr Ala Pro Ser Phe Lys Tyr Ile Cys Val His
 20 25 30

Thr Asp Thr His Ile Cys Val Cys Val Cys Ile Tyr Leu Ser Ser Val
 35 40 45
 5 Val Ser Lys Ser Ser Ala Glu Ala Asp Gly Val Leu Gln Pro Arg Arg
 50 55 60
 His Pro Ala Ser Leu Leu Ile Val Phe Ala Thr Ser Ile Ser Glu Ser
 65 70 75 80
 10 Ser Leu Leu Ile Phe Ser Phe Gln Lys Thr Glu Ala Lys Leu Ile Val
 85 90 95
 Phe Ala Val Ser Leu Ala Ala Lys Xaa
 15 100 105

- 20 (2) INFORMATION FOR SEQ ID NO: 458:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 70 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 458:

Met Leu Pro Pro Phe Ser Leu Val Tyr Thr His Phe Leu Val Ala Ser
 1 5 10 15
 30 Leu Leu Pro Val Ile Leu Ala Val Phe Pro Asp Ser Ala Gln Ile Val
 20 25 30
 Pro Leu Leu Lys Pro Ile Pro Arg Pro Gln Pro Glu Val Ile Phe Pro
 35 40 45
 35 Ser Ser Glu Leu Leu Glu Gln Leu Leu Ser Val Gln Phe Val Trp Gln
 50 55 60
 Ala His Thr Val Ala Xaa
 40 65 70

- 45 (2) INFORMATION FOR SEQ ID NO: 459:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 155 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 459:

Met Ala Leu Leu Leu Ser Val Leu Arg Val Leu Leu Gly Gly Phe Phe
 1 5 10 15
 55 Ala Leu Val Gly Leu Ala Lys Leu Ser Glu Glu Ile Ser Ala Pro Val
 20 25 30
 Ser Glu Arg Met Asn Ala Leu Phe Val Gln Phe Ala Glu Val Phe Pro
 35 40 45
 60

Leu Lys Val Phe Gly Tyr Gln Pro Asp Pro Leu Asn Tyr Gln Ile Ala
 50 55 60
 5 Val Gly Phe Leu Glu Leu Leu Ala Gly Leu Leu Leu Val Met Gly Pro
 65 70 75 80
 Pro Met Leu Gln Glu Ile Ser Asn Leu Phe Leu Ile Leu Leu Met Met
 85 90 95
 10 Gly Ala Ile Phe Thr Leu Ala Ala Leu Lys Glu Ser Leu Ser Thr Cys
 100 105 110
 Ile Pro Ala Ile Val Cys Leu Gly Phe Leu Leu Leu Leu Asn Val Gly
 115 120 125
 15 Gln Leu Leu Ala Gln Thr Lys Lys Val Val Arg Pro Thr Arg Lys Lys
 130 135 140
 Thr Leu Ser Thr Phe Lys Glu Ser Trp Lys Xaa
 145 150 155

(2) INFORMATION FOR SEQ ID NO: 460:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 332 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 460:

Met Lys Leu Gly Arg Ala Val Leu Gly Leu Leu Leu Leu Ala Pro Ser
 1 5 10 15
 35 Val Val Gln Ala Val Glu Pro Ile Ser Leu Gly Leu Ala Leu Ala Gly
 20 25 30
 Val Leu Thr Gly Tyr Ile Tyr Pro Arg Leu Tyr Cys Leu Phe Ala Glu
 35 40 45
 40 Cys Cys Gly Gln Lys Arg Ser Leu Ser Arg Glu Ala Leu Gln Lys Asp
 50 55 60
 45 Leu Asp Asp Asn Leu Phe Gly Gln His Leu Ala Lys Lys Ile Ile Leu
 65 70 75 80
 Asn Ala Val Phe Gly Phe Ile Asn Asn Pro Lys Pro Lys Lys Pro Leu
 85 90 95
 50 Thr Leu Ser Leu His Gly Trp Thr Gly Thr Gly Lys Asn Phe Val Ser
 100 105 110
 Lys Ile Ile Ala Glu Asn Ile Tyr Glu Gly Gly Leu Asn Ser Asp Tyr
 115 120 125
 55 Val His Leu Phe Val Ala Thr Leu His Phe Pro His Ala Ser Asn Ile
 130 135 140
 Thr Leu Tyr Lys Asp Gln Leu Gln Leu Trp Ile Arg Gly Asn Val Ser
 145 150 155 160

Ala Cys Ala Arg Ser Ile Phe Ile Phe Asp Glu Met Asp Lys Met His
165 170 175

5 Ala Gly Leu Ile Asp Ala Ile Lys Pro Phe Leu Asp Tyr Tyr Asp Leu
180 185 190

Val Asp Gly Val Ser Tyr Gln Lys Ala Met Phe Ile Phe Leu Ser Asn
195 200 205

10 Ala Gly Ala Glu Arg Ile Thr Asp Val Ala Leu Asp Phe Trp Arg Ser
210 215 220

Gly Lys Gln Arg Glu Asp Ile Lys Leu Lys Asp Ile Glu His Ala Leu
15 225 230 235 240

Ser Val Ser Val Phe Asn Asn Lys Asn Ser Gly Phe Trp His Ser Ser
245 250 255

20 Leu Ile Asp Arg Asn Leu Ile Asp Tyr Phe Val Pro Phe Leu Pro Leu
260 265 270

Glu Tyr Lys His Leu Lys Met Cys Ile Arg Val Glu Met Gln Ser Arg
275 280 285

25 Gly Tyr Glu Ile Asp Glu Asp Ile Val Ser Arg Val Ala Glu Glu Met
290 295 300

Thr Phe Phe Pro Lys Glu Arg Val Phe Ser Asp Lys Gly Cys Lys
30 305 310 315 320

Thr Val Phe Thr Lys Leu Asp Tyr Tyr Tyr Asp Asp
325 330

35

(2) INFORMATION FOR SEQ ID NO: 461:

- (i) SEQUENCE CHARACTERISTICS:
40 (A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 461:

45 Met Leu Lys Cys Ile
1 5

50 (2) INFORMATION FOR SEQ ID NO: 462:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
55 (D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 462:

Met Ile Leu Thr Leu Leu Ser Val Val Ser Thr Met Ala Ser
1 5 10
60

(2) INFORMATION FOR SEQ ID NO: 463:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 285 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 463:

10

Met Lys Leu His Pro Pro Pro Ser Pro Val Thr Gln Asp His Arg
1 5 10 15

15

Ser Lys Ser Ser His Ser Asn Trp Met Pro Arg Met Gly Ala Cys Ser
20 25 30

Met Ser Arg Thr Ser Ser Ser Gly Pro Pro Ser Leu Cys Lys Ser Thr
35 40 45

20

Ser Gly Arg Ser Cys Thr Arg Pro His Cys Trp Pro Ser Leu Pro Ala
50 55 60

Trp Val Ser Val Phe Thr Arg Thr Asn Thr Gly Ser Trp Cys Tyr Pro
65 70 75 80

25

Ala Trp Gly Gly Ala Phe Ser Arg Pro Trp Met Ser Ala Gln Ser Met
85 90 95

30

Cys Cys Ala Glu Arg Ser Val Leu Gln Val Ala Cys Arg Leu Leu Asp
100 105 110

Ala Leu Glu Phe Leu His Glu Asn Glu Tyr Val His Gly Asn Val Thr
115 120 125

35

Ala Glu Asn Ile Phe Val Asp Pro Glu Asp Gln Ser Gln Val Thr Leu
130 135 140

Ala Gly Tyr Gly Phe Ala Phe Arg Tyr Cys Pro Ser Gly Lys His Val
145 150 155 160

40

Ala Tyr Val Glu Gly Ser Arg Ser Pro His Glu Gly Asp Leu Glu Phe
165 170 175

45

Ile Ser Met Asp Leu His Lys Gly Cys Gly Pro Ser Arg Arg Xaa Asp
180 185 190

Leu Gln Ser Leu Gly Tyr Cys Met Leu Lys Trp Leu Tyr Gly Phe Leu
195 200 205

50

Pro Trp Thr Asn Cys Leu Pro Xaa Xaa Glu Asp Ile Met Lys Gln Lys
210 215 220

Gln Lys Phe Val Asp Lys Pro Gly Pro Phe Val Gly Pro Cys Gly His
225 230 235 240

55

Trp Ile Arg Pro Ser Glu Thr Leu Gln Lys Tyr Leu Lys Val Val Met
245 250 255

60

Ala Leu Thr Tyr Glu Glu Lys Pro Pro Tyr Ala Met Leu Arg Asn Asn
260 265 270

Leu Glu Ala Leu Leu Gln Asp Leu Arg Val Ser Pro Tyr
 275 280 285

5

(2) INFORMATION FOR SEQ ID NO: 464:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 80 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 464:

15

Met Thr Ser Pro Pro Pro His Gln Gly Trp Glu Gln Arg Gly Cys Gly
 1 5 10 15

Glu Ser Gln Val Pro Leu Ala Leu Ser Arg Val Phe Ser Thr Ser His
 20 25 30

20

Tyr Cys Leu Leu Leu Val Ala Asn Gln Ser Ile Phe Phe Pro Cys Leu
 35 40 45

25

Trp Ala Val Glu Arg Leu Leu Gly Val Arg Cys Thr Cys Pro Leu Ser
 50 55 60

Trp Gly Lys Arg Ile Ile Ser Glu His Cys Ser Ala Gln Ser Ser Xaa
 65 70 75 80

30

35

(2) INFORMATION FOR SEQ ID NO: 465:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 465:

Met His Thr Trp Tyr Asn Asp Arg Arg Gln Asn Cys His Cys Leu Leu
 1 5 10 15

45

Phe Phe Leu Ile Tyr Leu Arg Lys Ile Tyr Gln Val Val Pro His Val
 20 25 30

50

Pro Leu Leu Val Lys Cys Arg Gly Arg Leu Lys Gly Val Asn Ile
 35 40 45

55

(2) INFORMATION FOR SEQ ID NO: 466:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 96 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 466:

Met Glu Leu Val Leu Val Phe Leu Cys Ser Leu Leu Ala Pro Met Val
 1 5 10 15
 5 Leu Ala Ser Ala Ala Glu Lys Glu Lys Glu Met Asp Pro Phe His Tyr
 20 25 30
 Asp Tyr Gln Thr Leu Arg Ile Gly Gly Leu Val Phe Ala Val Val Leu
 35 40 45
 10 Phe Ser Val Gly Ile Leu Leu Ile Leu Ser Arg Arg Cys Lys Cys Ser
 50 55 60
 Phe Asn Gln Lys Pro Arg Ala Pro Gly Asp Glu Glu Ala Gln Val Glu
 15 65 70 75 80
 Asn Leu Ile Thr Ala Asn Ala Thr Glu Pro Gln Lys Ala Glu Asn Xaa
 85 90 95
 20

25 (2) INFORMATION FOR SEQ ID NO: 467:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 399 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 467:

Met Ala Ser Gly Ala Asp Ser Lys Gly Asp Asp Leu Ser Thr Ala Ile
 1 5 10 15
 35 Leu Lys Gln Lys Asn Arg Pro Asn Arg Leu Ile Val Asp Glu Ala Ile
 20 25 30
 Asn Glu Asp Asn Ser Val Val Ser Leu Ser Gln Pro Lys Met Asp Glu
 40 35 40 45
 Leu Gln Leu Phe Arg Gly Asp Thr Val Leu Leu Lys Gly Lys Lys Arg
 50 55 60
 45 Arg Glu Ala Val Cys Ile Val Leu Ser Asp Asp Thr Cys Ser Asp Glu
 65 70 75 80
 Lys Ile Arg Met Asn Arg Val Val Arg Asn Asn Leu Arg Val Arg Leu
 85 90 95
 50 Gly Asp Val Ile Ser Ile Gln Pro Cys Pro Asp Val Lys Tyr Gly Lys
 100 105 110
 Arg Ile His Val Leu Pro Ile Asp Asp Thr Val Glu Gly Ile Thr Gly
 55 115 120 125
 Asn Leu Phe Glu Val Tyr Leu Lys Pro Tyr Phe Leu Glu Ala Tyr Arg
 130 135 140
 60 Pro Ile Arg Lys Gly Asp Ile Phe Leu Val Arg Gly Gly Met Arg Ala

	145				150					155					160	
	Val	Glu	Phe	Lys	Val	Val	Glu	Thr	Asp	Pro	Ser	Pro	Tyr	Cys	Ile	Val
					165					170					175	
5	Ala	Pro	Asp	Thr	Val	Ile	His	Cys	Glu	Gly	Glu	Pro	Ile	Lys	Arg	Glu
				180					185					190		
10	Asp	Glu	Glu	Glu	Ser	Leu	Asn	Glu	Val	Gly	Tyr	Asp	Asp	Ile	Gly	Gly
		195						200					205			
	Cys	Arg	Lys	Gln	Leu	Ala	Gln	Ile	Lys	Glu	Met	Val	Glu	Leu	Pro	Leu
		210					215					220				
15	Arg	His	Pro	Ala	Leu	Phe	Lys	Ala	Ile	Gly	Val	Lys	Pro	Pro	Arg	Gly
	225					230					235					240
	Ile	Leu	Leu	Tyr	Gly	Pro	Pro	Gly	Thr	Gly	Lys	Thr	Leu	Ile	Ala	Arg
				245						250					255	
20	Ala	Val	Ala	Asn	Glu	Thr	Gly	Ala	Phe	Phe	Phe	Leu	Ile	Asn	Gly	Pro
			260						265					270		
	Glu	Ile	Met	Ser	Lys	Leu	Ala	Gly	Glu	Ser	Glu	Ser	Asn	Leu	Arg	Lys
25			275					280					285			
	Ala	Phe	Glu	Glu	Ala	Glu	Lys	Asn	Ala	Pro	Ala	Ile	Ile	Phe	Ile	Asp
		290					295					300				
30	Glu	Leu	Asp	Ala	Ile	Ala	Pro	Lys	Arg	Glu	Lys	Thr	His	Gly	Glu	Val
		305				310					315					320
	Glu	Arg	Arg	Ile	Val	Ser	Gln	Leu	Leu	Thr	Leu	Met	Asp	Gly	Leu	Lys
				325						330					335	
35	Gln	Arg	Ala	His	Val	Ile	Val	Met	Ala	Ala	Thr	Asn	Arg	Pro	Asn	Ser
			340						345					350		
	Ile	Asp	Pro	Ala	Leu	Arg	Arg	Phe	Gly	Arg	Phe	Asp	Arg	Glu	Val	Asp
40			355					360					365			
	Ile	Gly	Ile	Pro	Asp	Ala	Thr	Gly	Arg	Leu	Glu	Ile	Leu	Gln	Ile	His
		370					375					380				
45	Thr	Lys	Asn	Met	Lys	Leu	Ala	Asp	Asp	Val	Asp	Leu	Glu	Gln	Xaa	
		385				390					395					

50 (2) INFORMATION FOR SEQ ID NO: 468:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 468:

Leu

1

60

(2) INFORMATION FOR SEQ ID NO: 469:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 273 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 469:

10

Met Ala Ala Pro Lys Gly Ser Leu Trp Val Arg Thr Gln Leu Gly Leu
1 5 10 15

15

Pro Pro Leu Leu Leu Leu Thr Met Ala Leu Ala Gly Gly Ser Gly Thr
20 25 30

Ala Ser Ala Glu Ala Phe Asp Ser Val Leu Gly Asp Thr Ala Ser Cys
35 40 45

20

His Arg Ala Cys Gln Leu Thr Tyr Pro Leu His Thr Tyr Pro Lys Glu
50 55 60

25

Glu Glu Leu Tyr Ala Cys Gln Arg Gly Cys Arg Leu Phe Ser Ile Cys
65 70 75 80

Gln Phe Val Asp Asp Gly Ile Asp Leu Asn Arg Thr Lys Leu Glu Cys
85 90 95

30

Glu Ser Ala Cys Thr Glu Ala Tyr Ser Gln Ser Asp Glu Gln Tyr Ala
100 105 110

Cys His Leu Gly Cys Gln Asn Gln Leu Pro Phe Ala Glu Leu Arg Gln
115 120 125

35

Glu Gln Leu Met Ser Leu Met Pro Lys Met His Leu Leu Phe Pro Leu
130 135 140

40

Thr Leu Val Arg Ser Phe Trp Ser Asp Met Met Asp Ser Ala Gln Ser
145 150 155 160

Phe Ile Thr Ser Ser Trp Thr Phe Tyr Leu Gln Ala Asp Asp Gly Lys
165 170 175

45

Ile Val Ile Phe Xaa Ser Lys Pro Arg Asn Pro Arg Tyr Ala Pro His
180 185 190

Leu Glu Pro Gly Ala Leu Pro Asn Leu Xaa Xaa Xaa Ser Leu Ser Lys
195 200 205

50

Met Ser Xaa Xaa Ser Xaa Met Arg Asn Ser Gln Ala His Arg Asn Phe
210 215 220

55

Leu Glu Asp Gly Glu Ser Asp Gly Phe Leu Arg Cys Leu Ser Leu Asn
225 230 235 240

Ser Gly Trp Ile Leu Thr Thr Thr Leu Val Leu Ser Val Met Val Leu
245 250 255

60

Leu Trp Ile Cys Cys Ala Thr Cys Cys Tyr Thr Leu Leu Asp Ala Val
260 265 270

Xaa

5

(2) INFORMATION FOR SEQ ID NO: 470:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 192 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 470:

15

Met Met Val Leu Ser Leu Gly Ile Ile Leu Ala Ser Ala Ser Phe Ser
1 5 10 15

Pro Asn Phe Thr Gln Val Thr Ser Thr Leu Leu Asn Ser Ala Tyr Pro
20 25 30

20

Phe Ile Gly Pro Phe Phe Phe Ile Ile Ser Gly Ser Leu Ser Ile Ala
35 40 45

25

Thr Glu Lys Arg Leu Thr Lys Leu Leu Val His Ser Ser Leu Val Gly
50 55 60

Ser Ile Leu Ser Ala Leu Ser Ala Leu Val Gly Phe Ile Ile Leu Ser
65 70 75 80

30

Val Lys Gln Ala Thr Leu Asn Pro Ala Ser Leu Gln Cys Glu Leu Asp
85 90 95

Lys Asn Asn Ile Pro Thr Arg Ser Tyr Val Ser Tyr Phe Tyr His Asp
100 105 110

35

Ser Leu Tyr Thr Thr Asp Cys Tyr Thr Ala Lys Ala Ser Leu Ala Gly
115 120 125

40

Xaa Leu Ser Leu Met Leu Ile Cys Thr Leu Leu Glu Phe Cys Leu Ala
130 135 140

Val Leu Thr Ala Val Leu Arg Trp Lys Gln Ala Tyr Ser Asp Phe Pro
145 150 155 160

45

Gly Ser Val Leu Phe Leu Pro His Ser Tyr Ile Gly Asn Ser Gly Met
165 170 175

Ser Ser Lys Met Thr His Asp Cys Gly Tyr Glu Glu Leu Leu Thr Ser
180 185 190

50

55

(2) INFORMATION FOR SEQ ID NO: 471:

60

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 234 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 471:

5 Met Arg Lys Thr Arg Leu Trp Gly Leu Leu Trp Met Leu Phe Val Ser
 1 5 10 15
 Glu Leu Arg Ala Ala Thr Lys Leu Thr Glu Glu Lys Tyr Glu Leu Lys
 20 25 30
 10 Glu Gly Gln Thr Leu Asp Val Lys Cys Asp Tyr Thr Leu Glu Lys Phe
 35 40 45
 Ala Ser Ser Gln Lys Ala Trp Gln Ile Ile Arg Asp Gly Glu Met Pro
 50 55 60
 15 Lys Thr Leu Ala Cys Thr Glu Arg Pro Ser Lys Asn Ser His Pro Val
 65 70 75 80
 Gln Val Gly Arg Ile Ile Leu Glu Asp Tyr His Asp His Gly Leu Leu
 85 90 95
 Arg Val Arg Met Val Asn Leu Gln Val Glu Asp Ser Gly Leu Tyr Gln
 100 105 110
 25 Cys Val Ile Tyr Gln Pro Pro Lys Glu Pro His Met Leu Phe Asp Arg
 115 120 125
 Ile Arg Leu Val Val Thr Lys Gly Phe Ser Gly Thr Pro Gly Ser Asn
 130 135 140
 30 Glu Asn Ser Thr Gln Asn Val Tyr Lys Ile Pro Pro Thr Thr Thr Lys
 145 150 155 160
 Ala Leu Cys Pro Leu Tyr Thr Ser Pro Arg Thr Val Thr Gln Ala Pro
 165 170 175
 Pro Lys Ser Thr Ala Asp Val Ser Thr Pro Asp Ser Glu Ile Asn Leu
 180 185 190
 40 Thr Asn Val Thr Asp Ile Ile Arg Val Pro Val Phe Asn Ile Val Ile
 195 200 205
 Leu Leu Ala Gly Gly Phe Leu Ser Lys Ser Leu Val Phe Ser Val Leu
 210 215 220
 45 Phe Ala Val Thr Leu Arg Ser Phe Val Pro
 225 230

50

(2) INFORMATION FOR SEQ ID NO: 472:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 105 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 472:

60 Met Leu His Ile Leu Pro Leu Lys Ser Tyr Asp Phe Pro His Phe Ser
 1 5 10 15

Leu Met Gly Arg Tyr Arg Cys Ala Ser Leu Leu Phe Cys Phe Leu Leu
 20 25 30
 5 Leu Phe Phe Phe Phe Cys Ser Val Leu Trp Thr Phe Ser Asp Met His
 35 40 45
 Arg Ser Gly Glu Asp Gly Pro Trp Thr Pro Cys Val His His Leu Ala
 50 55 60
 10 Ala Ser Leu Ile Ser Tyr Gly Gln Pro Gly Phe Ile Cys Ile Ser Leu
 65 70 75 80
 Phe Ser Pro Val Leu Phe Ile Glu Asn Pro Arg His Tyr Ala Asn Ala
 15 85 90 95
 Thr Val Thr Thr Leu Gly Asp Trp Xaa
 100 105

20

(2) INFORMATION FOR SEQ ID NO: 473:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 473:

30 Met Val Phe Leu Lys Tyr Arg Phe Leu Phe Phe Leu Val Phe Leu Ala
 1 5 10 15
 Asn Cys Ile Tyr Ser Leu His Tyr Lys Pro Ser Leu Met Tyr Pro Lys
 35 20 25 30

40

(2) INFORMATION FOR SEQ ID NO: 474:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 571 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 474:

50 Met Ala Leu Ser Arg Gly Leu Pro Arg Glu Leu Ala Glu Ala Val Ala
 1 5 10 15
 Gly Gly Arg Val Leu Val Val Gly Ala Gly Gly Ile Gly Cys Glu Leu
 20 25 30
 55 Leu Lys Asn Leu Val Leu Thr Gly Phe Ser His Ile Asp Leu Ile Asp
 35 40 45
 Leu Asp Thr Ile Asp Val Ser Asn Leu Asn Arg Gln Phe Leu Phe Gln
 60 50 55 60

Lys Lys His Val Gly Arg Ser Lys Ala Gln Val Ala Lys Glu Ser Val
 65 70 75 80
 5 Leu Gln Phe Tyr Pro Lys Ala Asn Ile Val Ala Tyr His Asp Ser Ile
 85 90 95
 Met Asn Pro Asp Tyr Asn Val Glu Phe Phe Arg Gln Phe Ile Leu Val
 100 105 110
 10 Met Asn Ala Leu Asp Asn Arg Ala Ala Arg Asn His Val Asn Arg Met
 115 120 125
 Cys Leu Ala Ala Asp Val Pro Leu Ile Glu Ser Gly Thr Ala Gly Tyr
 130 135 140
 15 Leu Gly Gln Val Thr Thr Ile Lys Lys Gly Val Thr Glu Cys Tyr Glu
 145 150 155 160
 Cys His Pro Lys Pro Thr Gln Arg Thr Phe Pro Gly Cys Thr Ile Arg
 165 170 175
 20 Asn Thr Pro Ser Glu Pro Ile His Cys Ile Val Trp Ala Lys Tyr Leu
 180 185 190
 Phe Asn Gln Leu Phe Gly Glu Glu Asp Ala Asp Gln Glu Val Ser Pro
 195 200 205
 Asp Arg Ala Asp Pro Glu Ala Ala Trp Glu Pro Thr Glu Ala Glu Ala
 210 215 220
 30 Arg Ala Arg Ala Ser Asn Glu Asp Gly Asp Ile Lys Arg Ile Ser Thr
 225 230 235 240
 Lys Glu Trp Ala Lys Ser Thr Gly Tyr Asp Pro Val Lys Leu Phe Thr
 245 250 255
 35 Lys Leu Phe Lys Asp Asp Ile Arg Tyr Leu Leu Thr Met Asp Lys Leu
 260 265 270
 Trp Arg Lys Arg Lys Pro Pro Val Pro Leu Asp Trp Ala Glu Val Gln
 275 280 285
 Ser Gln Gly Glu Glu Thr Asn Ala Ser Asp Gln Gln Asn Glu Pro Gln
 290 295 300
 45 Leu Gly Leu Lys Asp Gln Gln Val Leu Asp Val Lys Ser Tyr Ala Arg
 305 310 315 320
 Leu Phe Ser Lys Ser Ile Glu Thr Leu Arg Val His Leu Ala Glu Lys
 325 330 335
 Gly Asp Gly Ala Glu Leu Ile Trp Asp Lys Asp Asp Pro Ser Ala Met
 340 345 350
 55 Asp Phe Val Thr Ser Ala Ala Asn Leu Arg Met His Ile Phe Ser Met
 355 360 365
 Asn Met Lys Ser Arg Phe Asp Ile Lys Ser Met Ala Gly Asn Ile Ile
 370 375 380
 60

Pro Ala Ile Ala Thr Thr Asn Ala Val Ile Ala Gly Leu Ile Val Leu
 385 390 395 400
 5 Glu Gly Leu Lys Ile Leu Ser Gly Lys Ile Asp Gln Cys Arg Thr Ile
 405 410 415
 Phe Leu Asn Lys Gln Pro Asn Pro Arg Lys Lys Leu Leu Val Pro Cys
 420 425 430
 10 Ala Leu Asp Pro Pro Asn Pro Asn Cys Tyr Val Cys Ala Ser Lys Pro
 435 440 445
 Glu Val Thr Val Arg Leu Asn Val His Lys Val Thr Val Leu Thr Leu
 450 455 460
 15 Gln Asp Lys Ile Val Lys Glu Lys Phe Ala Met Val Ala Pro Asp Val
 465 470 475 480
 Gln Ile Glu Asp Gly Lys Gly Thr Ile Leu Ile Ser Ser Glu Glu Gly
 485 490 495
 20 Glu Thr Glu Ala Asn Asn His Lys Lys Leu Ser Glu Phe Gly Ile Arg
 500 505 510
 25 Asn Gly Ser Arg Leu Gln Ala Asp Asp Phe Leu Gln Asp Tyr Thr Leu
 515 520 525
 Leu Ile Asn Ile Leu His Ser Glu Asp Leu Gly Lys Asp Val Glu Phe
 530 535 540
 30 Glu Val Val Gly Asp Ala Pro Glu Lys Val Gly Xaa Lys Gln Ala Glu
 545 550 555 560
 35 Asp Ala Ala Lys Ser Ile Thr Asn Gly Gln Xaa
 565 570

40 (2) INFORMATION FOR SEQ ID NO: 475:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 312 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 475:

Met Gln Val Val Thr Cys Leu Thr Arg Asp Ser Tyr Leu Thr His Cys
 1 5 10 15
 50 Phe Leu Gln His Leu Met Val Val Leu Ser Ser Leu Glu Arg Thr Pro
 20 25 30
 Ser Pro Glu Pro Val Asp Lys Asp Phe Tyr Ser Glu Phe Gly Asn Lys
 35 40 45
 55 Thr Thr Gly Lys Met Glu Asn Tyr Glu Leu Ile His Ser Ser Arg Val
 50 55 60
 60 Lys Phe Thr Tyr Pro Ser Glu Glu Glu Ile Gly Asp Leu Thr Phe Thr
 65 70 75 80

Val Ala Gln Lys Met Ala Glu Pro Glu Lys Ala Pro Ala Leu Ser Ile
 85 90 95
 5 Leu Leu Tyr Val Gln Ala Phe Gln Val Gly Met Pro Pro Pro Gly Cys
 100 105 110
 Cys Arg Gly Pro Leu Arg Pro Lys Thr Leu Leu Leu Thr Ser Ser Glu
 115 120 125
 10 Ile Phe Leu Leu Asp Glu Asp Cys Val His Tyr Pro Leu Pro Glu Phe
 130 135 140
 Ala Lys Glu Pro Pro Gln Arg Asp Arg Tyr Arg Leu Asp Asp Gly Arg
 15 145 150 155 160
 Arg Val Arg Asp Leu Asp Arg Val Leu Met Gly Tyr Gln Thr Tyr Pro
 165 170 175
 20 Gln Pro Ser Pro Ser Ser Ser Met Thr Cys Lys Val Met Thr Ser Trp
 180 185 190
 Ala Val Ser Pro Trp Thr Thr Leu Gly Arg Cys Gln Val Ala Arg Leu
 25 195 200 205
 Glu Pro Ala Arg Ala Val Lys Ser Ser Gly Arg Cys Leu Ser Pro Val
 210 215 220
 Leu Arg Ala Glu Arg Ser Ser Ser Arg Cys Trp Leu Ala Ser Gly Arg
 30 225 230 235 240
 Pro Cys Val Ala Val Ser Cys Leu Ser Ser Ser Pro Ala Ser Pro Gly
 245 250 255
 35 His Ser Gln Pro Val Val Ser Ser Leu Thr Pro Thr Gly Ala Gly Gln
 260 265 270
 Gln Ala Phe Val Phe Ser Lys Asn Val Leu Ser Ser Leu Trp Tyr Leu
 40 275 280 285
 Asn Leu Thr Val Leu Ala Glu Asn Val Asn Met Cys Val Cys Cys Val
 290 295 300
 Asn Ser Phe Ser Cys Trp Glu Xaa
 45 305 310

(2) INFORMATION FOR SEQ ID NO: 476:

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 329 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 476:

Met Ala Gln His His Leu Trp Ile Leu Leu Leu Cys Leu Gln Thr Trp
 1 5 10 15

60 Pro Glu Ala Ala Gly Lys Asp Ser Glu Ile Phe Thr Val Asn Gly Ile

	20	25	30
	Leu Gly Glu Ser Val Thr Phe Pro Val Asn Ile Gln Glu Pro Arg Gln		
	35	40	45
5	Val Lys Ile Ile Ala Trp Thr Ser Lys Thr Ser Val Ala Tyr Val Thr		
	50	55	60
10	Pro Gly Asp Ser Glu Thr Ala Pro Val Val Thr Val Thr His Arg Asn		
	65	70	75
	Tyr Tyr Glu Arg Ile His Ala Leu Gly Pro Asn Tyr Asn Leu Val Ile		
	85	90	95
15	Ser Asp Leu Arg Met Glu Asp Ala Gly Asp Tyr Lys Ala Asp Ile Asn		
	100	105	110
	Thr Gln Ala Asp Pro Tyr Thr Thr Thr Lys Arg Tyr Asn Leu Gln Ile		
	115	120	125
20	Tyr Arg Arg Leu Gly Lys Pro Lys Ile Thr Gln Ser Leu Met Ala Ser		
	130	135	140
	Val Asn Ser Thr Cys Asn Val Thr Leu Thr Cys Ser Val Glu Lys Glu		
25	145	150	155
	Glu Lys Asn Val Thr Tyr Asn Trp Ser Pro Leu Gly Glu Glu Gly Asn		
	165	170	175
30	Val Leu Gln Ile Phe Gln Thr Pro Glu Asp Gln Glu Leu Thr Tyr Thr		
	180	185	190
	Cys Thr Ala Gln Asn Pro Val Ser Asn Asn Ser Asp Ser Ile Ser Ala		
	195	200	205
35	Arg Gln Leu Cys Ala Asp Ile Ala Met Gly Phe Arg Thr His His Thr		
	210	215	220
	Gly Leu Leu Ser Val Leu Ala Met Phe Phe Leu Leu Val Leu Ile Leu		
40	225	230	235
	Ser Ser Val Phe Leu Phe Arg Leu Phe Lys Arg Arg Gln Asp Ala Ala		
	245	250	255
45	Ser Lys Lys Thr Ile Tyr Thr Tyr Ile Met Ala Ser Arg Asn Thr Gln		
	260	265	270
	Pro Ala Glu Ser Arg Ile Tyr Asp Glu Ile Leu Gln Ser Lys Val Leu		
	275	280	285
50	Pro Ser Lys Glu Glu Pro Val Asn Thr Val Tyr Ser Glu Val Gln Phe		
	290	295	300
	Ala Asp Lys Met Gly Lys Ala Ser Thr Gln Asp Ser Lys Pro Pro Gly		
55	305	310	315
	Thr Ser Ser Tyr Glu Ile Val Ile Xaa		
	325		
60			

(2) INFORMATION FOR SEQ ID NO: 477:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 178 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 477:

5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 178 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 477:
 10 Met Lys Leu Gln Cys Val Ser Leu Trp Leu Leu Gly Thr Ile Leu Ile
 1 5 10 15
 Leu Cys Ser Val Asp Asn His Gly Leu Arg Arg Cys Leu Ile Ser Thr
 20 25 30
 15 Asp Met His His Ile Glu Glu Ser Phe Gln Glu Ile Lys Arg Ala Ile
 35 40 45
 Gln Ala Lys Asp Thr Phe Pro Asn Val Thr Ile Leu Ser Thr Leu Glu
 20 50 55 60
 Thr Leu Gln Ile Ile Lys Pro Leu Asp Val Cys Cys Val Thr Lys Asn
 65 70 75 80
 25 Leu Leu Ala Phe Tyr Val Asp Arg Val Phe Lys Asp His Gln Glu Pro
 85 90 95
 Asn Pro Lys Ile Leu Arg Lys Ile Ser Ser Ile Ala Asn Ser Phe Leu
 100 105 110
 30 Tyr Met Gln Lys Thr Leu Arg Gln Cys Gln Glu Gln Arg Gln Cys His
 115 120 125
 Cys Arg Gln Glu Ala Thr Asn Ala Thr Arg Val Ile His Asp Asn Tyr
 35 130 135 140
 Asp Gln Leu Glu Val His Ala Ala Ala Ile Lys Ser Leu Gly Glu Leu
 145 150 155 160
 40 Asp Val Phe Leu Ala Trp Ile Asn Lys Asn His Glu Val Met Ser Ser
 165 170 175
 Ala Xaa

(2) INFORMATION FOR SEQ ID NO: 478:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 52 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 478:

50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 52 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 478:
 55 Asp Thr Ala Ile Arg Val Ala Leu Ala Val Ala Val Leu Lys Thr Val
 1 5 10 15
 Ile Leu Gly Leu Leu Cys Leu Leu Leu Cys Gly Gly Gly Glu Gly Lys
 60 20 25 30

Val Ala Gly Arg Gln Ala Val Thr Ser Asp Gln Gln Ser Val Gly Arg
 35 40 45

5 Arg Asp Val Tyr
 50

10 (2) INFORMATION FOR SEQ ID NO: 479:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 62 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 479:

Met Gln Lys Lys Asn Ser Leu Phe Phe Phe Phe Ala Phe Tyr Tyr Glu
 1 5 10 15

20

Asn Lys Thr Asn Ala Pro Gly Glu Gly Ser Met Ile Thr Arg Asn Ile
 20 25 30

25

Lys Glu Tyr Phe Leu Pro Phe Leu Phe Cys Cys Val Glu Ala Ser Ile
 35 40 45

Ala Ile Asn Lys Leu Asn Tyr Leu His Trp Thr His Phe Gln
 50 55 60

30

(2) INFORMATION FOR SEQ ID NO: 480:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 480:

40

Met Pro Gly Leu Ser Leu Ile Leu Thr Val Thr Leu Leu Ala Val Ser
 1 5 10 15

Asp Ser Ala Ala Thr Cys Ile Val Ala Lys Gly
 20 25

45

(2) INFORMATION FOR SEQ ID NO: 481:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 339 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 481:

55

Met Ser Gly Pro Asp Val Glu Thr Pro Ser Ala Ile Gln Ile Cys Arg
 1 5 10 15

60

Ile Met Arg Pro Asp Asp Ala Asn Val Ala Gly Asn Val His Gly Gly
 20 25 30

Thr Ile Leu Lys Met Ile Glu Glu Ala Gly Ala Ile Ile Ser Thr Arg
 35 40 45
 5 His Cys Asn Ser Gln Asn Gly Glu Arg Cys Val Ala Ala Leu Ala Arg
 50 55 60
 Val Glu Arg Thr Asp Phe Leu Ser Pro Met Cys Ile Gly Glu Val Ala
 65 70 75 80
 10 His Val Ser Ala Glu Ile Thr Tyr Thr Ser Lys His Ser Val Glu Val
 85 90 95
 Gln Val Asn Val Met Ser Glu Asn Ile Leu Thr Gly Ala Lys Lys Leu
 15 100 105 110
 Thr Asn Lys Ala Thr Leu Trp Tyr Val Pro Leu Ser Leu Lys Asn Val
 115 120 125
 20 Asp Lys Val Leu Glu Val Pro Pro Val Val Tyr Ser Arg Xaa Glu Gln
 130 135 140
 Glu Glu Glu Gly Arg Lys Arg Tyr Glu Ala Gln Lys Leu Glu Arg Met
 145 150 155 160
 25 Glu Thr Lys Trp Arg Asn Gly Asp Ile Val Gln Pro Val Leu Asn Pro
 165 170 175
 Glu Pro Asn Thr Val Ser Tyr Ser Gln Ser Ser Leu Ile His Leu Val
 180 185 190
 Gly Pro Ser Asp Cys Thr Leu His Gly Phe Val His Gly Gly Val Thr
 195 200 205
 35 Met Lys Leu Met Asp Glu Val Ala Gly Ile Val Ala Ala Arg His Cys
 210 215 220
 Lys Thr Asn Ile Val Thr Ala Ser Val Asp Ala Ile Asn Phe His Asp
 225 230 235 240
 40 Lys Ile Arg Lys Gly Cys Val Ile Thr Ile Ser Gly Arg Met Thr Phe
 245 250 255
 Thr Ser Asn Lys Ser Met Glu Ile Glu Val Leu Val Asp Ala Asp Pro
 260 265 270
 Val Val Asp Ser Ser Gln Lys Arg Tyr Arg Ala Ala Ser Ala Phe Phe
 275 280 285
 50 Thr Tyr Val Ser Leu Ser Gln Glu Gly Arg Ser Leu Pro Val Pro Gln
 290 295 300
 Leu Val Pro Glu Thr Glu Asp Glu Lys Lys Arg Phe Glu Glu Gly Lys
 305 310 315 320
 55 Gly Arg Tyr Leu Gln Met Lys Ala Lys Xaa Gln Gly His Ala Xaa Xaa
 325 330 335
 60 Gln Pro Xaa

(2) INFORMATION FOR SEQ ID NO: 482:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 482:

Met Leu Asn Ser Asn Ile Asn Asp Leu Leu Met Val Thr Tyr Leu Ala
1 5 10 15

Asn Leu Thr Gln Ser Gln Ile Ala Leu Asn Glu Lys Leu Val Asn Leu
20 25 30

20

(2) INFORMATION FOR SEQ ID NO: 483:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 48 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 483:

Met Arg Glu Thr Ser Ile Arg Val Leu Leu Met Leu Pro Ala Leu Glu
1 5 10 15

35

Ser Thr Ser Gly Leu Ser Ala Phe Met Gly Leu Gly Thr Arg Ile Gly
20 25 30

Cys Phe Lys Thr Ile Thr Cys Trp Pro Thr Ser Leu Thr Gln Arg Xaa
35 40 45

40

45

(2) INFORMATION FOR SEQ ID NO: 484:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 38 amino acids

(B) TYPE: amino acid

50

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 484:

Met Tyr Met Tyr Ser Leu Asn Val Phe Leu Ser Phe Ile Phe Leu Ala
1 5 10 15

55

Leu Val Phe Lys Cys Val His Val Cys Gln Gly Ala Asn Ala Phe Leu
20 25 30

60

Phe Leu Lys Leu Val Phe
35

(2) INFORMATION FOR SEQ ID NO: 485:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 61 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 485:

Met Gly Leu Arg Leu Ile Cys Leu Glu Leu Thr Met Val Lys Ala Leu
 1 5 10 15

15

Val Cys Glu Met Phe Leu Phe Phe Leu Met Thr Gln Lys Leu Ile Trp
 20 25 30

Gln Glu Cys Thr Glu Lys Phe Ala Lys Leu Leu Val Gln Leu Ile Ser
 35 40 45

20

Leu Val Phe Ala Trp Glu Phe Phe Ser Glu Asp Thr Pro
 50 55 60

25

(2) INFORMATION FOR SEQ ID NO: 486:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 346 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 486:

Met Leu Ala Ala Arg Leu Val Cys Leu Arg Thr Leu Pro Ser Arg Val
 35 1 5 10 15

Phe His Pro Ala Phe Thr Lys Ala Ser Pro Val Val Lys Asn Ser Ile
 20 25 30

40

Thr Lys Asn Gln Trp Leu Leu Thr Pro Ser Arg Glu Tyr Ala Thr Lys
 35 40 45

Thr Arg Ile Gly Ile Arg Arg Gly Arg Thr Gly Gln Glu Leu Lys Glu
 50 55 60

45

Ala Ala Leu Glu Pro Ser Met Glu Lys Ile Phe Lys Ile Asp Gln Met
 65 70 75 80

50

Gly Arg Trp Phe Val Ala Gly Gly Ala Ala Val Gly Leu Gly Ala Leu
 85 90 95

Cys Tyr Tyr Gly Leu Gly Leu Ser Asn Glu Ile Gly Ala Ile Glu Lys
 100 105 110

55

Ala Val Ile Trp Pro Gln Tyr Val Lys Asp Arg Ile His Ser Thr Tyr
 115 120 125

Met Tyr Leu Ala Gly Ser Ile Gly Leu Thr Ala Leu Ser Ala Ile Ala
 130 135 140

60

Ile Ser Arg Thr Pro Val Leu Met Asn Phe Met Met Arg Gly Ser Trp
 145 150 155 160
 5 Val Thr Ile Gly Val Thr Phe Ala Ala Met Val Gly Ala Gly Met Leu
 165 170 175
 Val Arg Ser Ile Pro Tyr Asp Gln Ser Pro Gly Pro Lys His Leu Ala
 180 185 190
 10 Trp Leu Leu His Ser Gly Val Met Gly Ala Val Val Ala Pro Leu Thr
 195 200 205
 Ile Leu Gly Gly Pro Leu Leu Ile Arg Ala Ala Trp Tyr Thr Ala Gly
 210 215 220
 15 Ile Val Gly Gly Leu Ser Thr Val Ala Met Cys Ala Pro Ser Glu Lys
 225 230 235 240
 Phe Leu Asn Met Gly Ala Pro Leu Gly Val Gly Leu Gly Leu Val Phe
 245 250 255
 20 Val Ser Ser Leu Gly Ser Met Phe Leu Pro Pro Thr Thr Val Ala Gly
 260 265 270
 25 Ala Thr Leu Tyr Ser Val Ala Met Tyr Gly Gly Leu Val Leu Phe Ser
 275 280 285
 Met Phe Leu Leu Tyr Asp Thr Gln Lys Val Ile Lys Arg Ala Glu Val
 290 295 300
 30 Ser Pro Met Tyr Gly Val Gln Lys Tyr Asp Pro Ile Asn Ser Met Leu
 305 310 315 320
 Ser Ile Tyr Met Asp Thr Leu Asn Ile Phe Met Arg Val Ala Thr Met
 325 330 335
 35 Leu Ala Thr Gly Gly Asn Arg Lys Lys Xaa
 340 345

40

(2) INFORMATION FOR SEQ ID NO: 487:

- 45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 237 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 487:

50 Met Glu Glu Val Leu Leu Leu Gly Leu Lys Asp Arg Glu Gly Tyr Thr
 1 5 10 15
 Ser Phe Trp Asn Asp Cys Ile Ser Ser Gly Leu Arg Gly Cys Met Leu
 20 25 30
 55 Ile Glu Leu Ala Leu Arg Gly Arg Leu Gln Leu Glu Ala Cys Gly Met
 35 40 45
 Arg Arg Lys Ser Leu Leu Thr Arg Lys Val Ile Cys Lys Ser Asp Ala
 50 55 60

Pro Thr Gly Asp Val Leu Leu Asp Glu Ala Leu Lys His Val Lys Glu
 65 70 75 80
 5 Thr Gln Pro Pro Glu Thr Val Gln Asn Trp Ile Glu Leu Leu Ser Gly
 85 90 95
 Glu Thr Trp Asn Pro Leu Lys Leu His Tyr Gln Leu Arg Asn Val Arg
 100 105 110
 10 Glu Arg Leu Ala Lys Asn Leu Val Glu Lys Gly Val Leu Thr Thr Glu
 115 120 125
 Lys Gln Asn Phe Leu Leu Phe Asp Met Thr Thr His Pro Leu Thr Asn
 15 130 135 140
 Asn Asn Ile Lys Gln Arg Leu Ile Lys Lys Val Gln Glu Ala Val Leu
 145 150 155 160
 20 Asp Lys Trp Val Asn Asp Pro His Arg Met Asp Arg Arg Leu Leu Ala
 165 170 175
 Leu Ile Tyr Leu Ala His Ala Ser Asp Val Leu Glu Asn Ala Phe Ala
 180 185 190
 25 Pro Leu Leu Asp Glu Gln Tyr Asp Leu Ala Thr Lys Arg Val Arg Gln
 195 200 205
 Leu Leu Asp Leu Asp Pro Glu Val Glu Cys Leu Lys Ala Asn Thr Asn
 30 210 215 220
 Glu Val Leu Trp Ala Val Val Ala Ala Phe Thr Lys Xaa
 225 230 235
 35

(2) INFORMATION FOR SEQ ID NO: 488:

(i) SEQUENCE CHARACTERISTICS:
 40 (A) LENGTH: 200 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 488:
 45 Met Ala Gln Arg Met Val Trp Val Asp Leu Glu Met Thr Gly Leu Asp
 1 5 10 15
 Ile Glu Lys Asp Gln Ile Ile Glu Met Ala Cys Leu Ile Thr Asp Ser
 20 25 30
 50 Asp Leu Asn Ile Leu Ala Glu Gly Pro Asn Leu Ile Ile Lys Gln Pro
 35 40 45
 Asp Glu Leu Leu Asp Ser Met Ser Asp Trp Cys Lys Glu His His Gly
 55 50 55 60
 Lys Ser Gly Leu Thr Lys Ala Val Lys Glu Ser Thr Ile Thr Leu Gln
 65 70 75 80
 60 Gln Ala Glu Tyr Glu Phe Leu Ser Phe Val Arg Gln Gln Thr Pro Pro

650

85 90 95

Gly Leu Cys Pro Leu Ala Gly Asn Ser Val His Glu Asp Lys Lys Phe
100 105 110

5 Leu Asp Lys Tyr Met Pro Gln Phe Met Lys His Leu His Tyr Arg Ile
115 120 125

10 Ile Asp Val Ser Thr Val Lys Glu Leu Cys Arg Arg Trp Tyr Pro Glu
130 135 140

Glu Tyr Glu Phe Ala Pro Lys Lys Ala Ala Ser His Arg Ala Leu Asp
145 150 155 160

15 Asp Ile Ser Glu Ser Ile Lys Glu Leu Gln Phe Tyr Arg Asn Asn Ile
165 170 175

Phe Lys Lys Lys Ile Asp Glu Lys Lys Arg Lys Ile Ile Glu Asn Gly
180 185 190

20 Glu Asn Glu Lys Thr Val Ser Xaa
195 200

25

(2) INFORMATION FOR SEQ ID NO: 489:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 351 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 489:

35 Met Ala Thr Thr Ala Ala Pro Ala Gly Gly Ala Arg Asn Gly Ala Gly
1 5 10 15

Pro Glu Trp Gly Gly Phe Glu Glu Asn Ile Gln Gly Gly Gly Ser Ala
20 25 30

40 Val Ile Asp Met Glu Asn Met Asp Asp Thr Ser Gly Ser Ser Phe Glu
35 40 45

Asp Met Gly Glu Leu His Gln Arg Leu Arg Glu Glu Glu Val Asp Ala
50 55 60

45 Asp Ala Ala Asp Ala Ala Ala Glu Glu Glu Asp Gly Glu Phe Leu
65 70 75 80

50 Gly Met Lys Gly Phe Lys Gly Gln Leu Ser Arg Gln Val Ala Asp Gln
85 90 95

Met Trp Gln Ala Gly Lys Arg Gln Ala Ser Arg Ala Phe Ser Leu Tyr
100 105 110

55 Ala Asn Ile Asp Ile Leu Arg Pro Tyr Phe Asp Val Glu Pro Ala Gln
115 120 125

Val Arg Thr Gly Leu Leu Glu Ser Met Ile Pro Ile Lys Met Val Asn
130 135 140

60

Phe Pro Gln Lys Ile Ala Gly Glu Leu Tyr Gly Pro Leu Met Leu Val
 145 150 155 160
 5 Phe Thr Leu Val Ala Ile Leu Leu His Gly Met Lys Thr Ser Asp Thr
 165 170 175
 Ile Ile Arg Glu Gly Thr Leu Met Gly Thr Ala Ile Gly Thr Cys Phe
 180 185 190
 10 Gly Tyr Trp Leu Gly Val Ser Ser Phe Ile Tyr Phe Leu Ala Tyr Leu
 195 200 205
 Cys Asn Ala Gln Ile Thr Met Leu Gln Met Leu Ala Leu Leu Gly Tyr
 210 215 220
 15 Gly Leu Phe Gly His Cys Ile Val Leu Phe Ile Thr Tyr Asn Ile His
 225 230 235 240
 Leu His Ala Leu Phe Tyr Leu Phe Trp Leu Leu Val Gly Gly Leu Ser
 245 250 255
 20 Thr Leu Arg Met Val Ala Val Leu Val Ser Arg Thr Val Gly Pro Thr
 260 265 270
 25 Gln Arg Leu Leu Leu Cys Gly Thr Leu Ala Ala Leu His Met Leu Phe
 275 280 285
 Leu Leu Tyr Leu His Phe Ala Tyr His Lys Val Val Glu Gly Ile Leu
 290 295 300
 30 Asp Thr Leu Glu Gly Pro Asn Ile Pro Pro Ile Gln Arg Val Pro Arg
 305 310 315 320
 Asp Ile Pro Ala Met Leu Pro Ala Ala Arg Leu Pro Thr Thr Val Leu
 325 330 335
 35 Asn Ala Thr Ala Lys Ala Val Ala Val Thr Leu Gln Ser His Xaa
 340 345 350

40

(2) INFORMATION FOR SEQ ID NO: 490:

- (i) SEQUENCE CHARACTERISTICS:
 45 (A) LENGTH: 265 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 490:

50 Met Arg Gly Ser Arg Gly Gly Trp Ala Gly Glu Met Ala Ala Ser Gly
 1 5 10 15
 Glu Ser Gly Thr Ser Gly Gly Gly Gly Ser Thr Glu Glu Ala Phe Met
 20 25 30
 55 Thr Phe Tyr Ser Glu Val Lys Gln Ile Glu Lys Arg Asp Ser Val Leu
 35 40 45
 60 Thr Ser Lys Asn Gln Ile Glu Arg Leu Thr Arg Pro Gly Ser Ser Tyr
 50 55 60

Phe Asn Leu Asn Pro Phe Glu Val Leu Gln Ile Asp Pro Glu Val Thr
 65 70 75 80
 5 Asp Glu Glu Ile Lys Lys Arg Phe Arg Gln Leu Ser Ile Leu Val His
 85 90 95
 Pro Asp Lys Asn Gln Asp Asp Ala Asp Arg Ala Gln Lys Ala Phe Glu
 100 105 110
 10 Ala Val Asp Lys Ala Tyr Lys Leu Leu Asp Gln Glu Gln Lys Lys
 115 120 125
 Arg Ala Leu Asp Val Ile Gln Ala Gly Lys Glu Tyr Val Glu His Thr
 130 135 140
 15 Val Lys Glu Arg Lys Lys Gln Leu Lys Lys Glu Gly Lys Pro Thr Ile
 145 150 155 160
 20 Val Glu Glu Asp Asp Pro Glu Leu Phe Lys Gln Ala Val Tyr Lys Gln
 165 170 175
 Thr Met Lys Leu Phe Ala Glu Leu Glu Ile Lys Arg Lys Glu Arg Glu
 180 185 190
 25 Ala Lys Glu Met His Glu Arg Lys Arg Gln Arg Glu Glu Glu Ile Glu
 195 200 205
 Ala Gln Glu Lys Ala Lys Arg Glu Arg Glu Trp Gln Lys Asn Phe Glu
 210 215 220
 Glu Ser Arg Asp Gly Arg Val Asp Ser Trp Arg Asn Phe Gln Ala Asn
 225 230 235 240
 35 Thr Lys Gly Lys Lys Glu Lys Lys Asn Arg Thr Phe Leu Arg Pro Pro
 245 250 255
 Lys Val Lys Met Glu Gln Arg Glu Xaa
 260 265
 40

(2) INFORMATION FOR SEQ ID NO: 491:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 491:

50 Asp Ser Met Pro Thr Cys Pro Leu Xaa Ala Ser Leu Glu Cys Gly Pro
 1 5 10 15
 Leu Leu Pro Val Arg Leu Cys Cys Leu
 20 25
 55

(2) INFORMATION FOR SEQ ID NO: 492:

60

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 159 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 492:

Met Asn Glu Tyr Arg Val Pro Glu Leu Asn Val Gln Asn Gly Val Leu
 1 5 10 15

10 Lys Ser Leu Ser Phe Leu Phe Glu Tyr Ile Gly Glu Met Gly Lys Asp
 20 25 30

Tyr Ile Tyr Ala Val Thr Pro Leu Leu Glu Asp Ala Leu Met Asp Arg
 35 40 45

15 Asp Leu Val His Arg Gln Thr Ala Ser Ala Val Val Gln His Met Ser
 50 55 60

Leu Gly Val Tyr Gly Phe Gly Cys Glu Asp Ser Leu Asn His Leu Leu
 20 65 70 75 80

Asn Tyr Val Trp Pro Asn Val Phe Glu Thr Ser Pro His Val Ile Gln
 85 90 95

25 Ala Val Met Gly Ala Leu Glu Gly Leu Arg Val Ala Ile Gly Pro Cys
 100 105 110

Arg Met Leu Gln Tyr Cys Leu Gln Gly Leu Phe His Pro Ala Arg Lys
 115 120 125

30 Val Arg Asp Val Tyr Trp Lys Ile Tyr Asn Ser Ile Tyr Ile Gly Ser
 130 135 140

Gln Asp Ala Leu Ile Ala His Tyr Pro Arg Ile Tyr Gln Arg Xaa
 35 145 150 155

40 (2) INFORMATION FOR SEQ ID NO: 493:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 279 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 493:

Met Ile Ser Asp Asn Ser Ala Glu Asn Ile Ala Leu Val Thr Ser Met
 1 5 10 15

50 Tyr Asp Gly Leu Leu Gln Ala Gly Ala Arg Leu Cys Pro Thr Val Gln
 20 25 30

Leu Glu Asp Ile Arg Asn Leu Gln Asp Leu Thr Pro Leu Lys Leu Ala
 35 40 45

55 Ala Lys Glu Gly Lys Ile Glu Ile Phe Arg His Ile Leu Gln Arg Glu
 50 55 60

Phe Ser Gly Leu Ser His Leu Ser Arg Lys Phe Thr Glu Trp Cys Tyr
 60 65 70 75 80

Gly Pro Val Arg Val Ser Leu Tyr Asp Leu Ala Ser Val Asp Ser Cys
 85 90 95
 5 Glu Glu Asn Ser Val Leu Glu Ile Ile Ala Phe His Cys Lys Ser Pro
 100 105 110
 His Arg His Arg Met Val Val Leu Glu Pro Leu Asn Lys Leu Leu Gln
 115 120 125
 10 Ala Lys Trp Asp Leu Leu Ile Pro Lys Phe Phe Leu Asn Phe Leu Cys
 130 135 140
 Asn Leu Ile Tyr Met Phe Ile Phe Thr Ala Val Ala Tyr His Gln Pro
 15 145 150 155 160
 Thr Leu Lys Lys Gln Ala Ala Pro His Leu Lys Ala Glu Val Gly Asn
 165 170 175
 20 Ser Met Leu Leu Thr Gly His Ile Leu Ile Leu Leu Gly Gly Ile Tyr
 180 185 190
 Leu Leu Val Gly Gln Leu Trp Tyr Phe Trp Arg Arg His Val Phe Ile
 195 200 205
 25 Trp Ile Ser Phe Ile Asp Ser Tyr Phe Glu Ile Leu Phe Leu Phe Gln
 210 215 220
 Ala Leu Leu Thr Val Val Ser Gln Val Leu Cys Phe Leu Xaa Ile Glu
 30 225 230 235 240
 Trp Tyr Leu Pro Leu Leu Val Ser Ala Leu Val Leu Gly Trp Leu Asn
 245 250 255
 35 Leu Leu Tyr Tyr Thr Arg Gly Phe Gln His Thr Gly Ile Tyr Ser Val
 260 265 270
 Met Ile Gln Lys Pro Trp Xaa
 275
 40

(2) INFORMATION FOR SEQ ID NO: 494:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 193 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 494:
 50 Met Ile Arg Cys Gly Leu Ala Cys Glu Arg Cys Arg Trp Ile Leu Pro
 1 5 10 15
 Leu Leu Leu Leu Ser Ala Ile Ala Phe Asp Ile Ile Ala Leu Ala Gly
 55 20 25 30
 Arg Gly Trp Leu Gln Ser Ser Asp His Gly Gln Thr Ser Ser Leu Trp
 35 40 45
 60 Trp Lys Cys Ser Gln Glu Gly Gly Ser Gly Ser Tyr Glu Glu Gly

[illegible]

35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 205 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

40	Met	Ala	Ala	Gly	Asp	Gln	Val	Phe	Ser	Gly	Ala	Gly	His	Val	Xaa	Glu
	1				5					10					15	
	His	Val	Ala	Gly	Gly	Arg	His	Ala	Trp	Leu	Leu	Thr	Trp	Gln	Ser	Ala
				20					25					30		
45	Cys	Pro	Ala	Asn	Arg	Leu	Ser	Leu	Val	Pro	Leu	Val	Pro	Ser	Ala	Ser
			35					40					45			
	Met	Thr	Arg	Leu	Met	Arg	Xaa	Arg	Thr	Ala	Ser	Gly	Ser	Ser	Val	Ile
50		50					55					60				
	Leu	Trp	Met	Ala	Pro	Ala	Ala	Ala	Pro	Thr	Pro	Ala	Arg	Ala	Pro	Glu
	65				70						75				80	
55	Ala	Ala	Pro	Thr	Pro	Ala	Arg	Ala	Pro	Ala	Ala	Ala	Arg	Thr	Pro	Ala
					85					90					95	
	Arg	Gly	Pro	Thr	Trp	Thr	Ser	Pro	Pro	Thr	Arg	Val	Leu	Leu	Gly	Thr
				100					105					110		

Xaa Pro Gly Pro Ser Pro Trp Arg Ser Pro Ala Arg Arg Pro Ala Gln
 115 120 125
 5 Leu Pro Pro Pro Asp Ser Asp Leu Cys Ser Gly Pro Leu Leu Pro Gly
 130 135 140
 Pro Phe Ser Pro Pro Ala Cys His Thr Ala Pro Asn Ser Val Leu Ile
 145 150 155 160
 10 Gln Ser Leu Phe Cys Lys Ser Glu Leu Trp Trp Arg Gln Met Arg Ser
 165 170 175
 Ile Thr Trp Val Pro Ser Pro Lys Ala Gly Trp Arg Trp Thr Lys Gly
 180 185 190
 15 Arg Lys Gln Ala Ser Pro His Arg Ile Leu Phe His Xaa
 195 200 205
 20
 (2) INFORMATION FOR SEQ ID NO: 496:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 147 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 496:
 25 Met Ala Leu Thr Leu Leu Pro Ser Val Ser Arg Leu Pro Gly Glu Arg
 1 5 10 15
 Met Ala Ala Ser Gly Leu Pro Tyr Val Leu His His Lys Ser Ser Leu
 20 25 30
 35 Met Lys Val Ile Phe Phe Pro Tyr Pro Val Leu Pro Leu Pro Ala Pro
 35 40 45
 Asn Gly Thr Trp Val Pro Arg Leu Val Leu Gly Leu Gly Ser Gly Asp
 50 55 60
 40 Gln Val His Tyr Leu Pro Ile Ser Ser Ser Ile Val Asn Tyr Gly Thr
 65 70 75 80
 Ser Val Ser Gly Lys Ser Trp Val Phe Leu Val Tyr Pro Leu His Pro
 85 90 95
 45 Thr Pro Thr Trp Ser Thr Arg Cys Phe Gln Val Trp Asp Leu Leu Ser
 100 105 110
 50 Val Glu Leu Pro Asp Lys Gly Glu Gly Asn Thr Arg Arg Ala Ser Gly
 115 120 125
 Val Pro Gly Leu Ser Gln Leu Pro Thr Ser His Lys Pro Ile Lys Gln
 130 135 140
 55 Glu Tyr Xaa
 145
 60

(2) INFORMATION FOR SEQ ID NO: 497:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 64 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 497:

5
10 Met Val Trp Val Leu Trp Ser Ala Pro Ser Leu Ala Pro Pro Trp Val
1 5 10 15
Gly Pro Cys Trp Pro Ser Thr Gly Asn Cys Cys Leu Cys Glu Val Gly
20 25 30
15 Ala Ala Leu Pro Pro Arg Gly Pro Ser Leu Ser Asp Cys Leu Gly Leu
35 40 45
Pro Pro Trp Thr Pro Trp Gly Pro Ala Trp Thr Leu Ala Gln Ser Xaa
50 55 60
20

25

(2) INFORMATION FOR SEQ ID NO: 498:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 94 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 498:

30
35 Met Ser Thr Gly Ala Leu Asn Thr Ser Pro Pro Ala Ser Asn Arg Leu
1 5 10 15
Glu Ser Thr Leu Asn Glu Tyr Leu Ile Gln Pro Gln Leu His Cys Ser
20 25 30
40 Ser Val Gln Arg Leu Thr Leu Lys Trp Gly Cys Ser Ser Leu Gln Arg
35 40 45
Asp Gly Gln Ala Val Pro Trp Gly Leu Trp Gln Arg Ala Tyr Pro Ser
50 55 60
45 Leu Leu Pro Thr Leu Pro Ser Asp Leu Leu Arg Pro His Ala Val Thr
65 70 75 80
50 Pro Ser Val Ser Val Ser Val His Thr Cys Glu Ser Ser Xaa
85 90

55

(2) INFORMATION FOR SEQ ID NO: 499:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 499:

60

Met Phe Leu Ile Phe Val Tyr Phe Leu Lys Xaa Leu Phe Ser Ser Ser
1 5 10 15

5 Leu Pro Phe Leu Trp Leu
20

10 (2) INFORMATION FOR SEQ ID NO: 500:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 500:

Arg Gly Gly Leu Cys Pro Leu Leu Val Pro Gly Pro Leu Ala Arg Gln
1 5 10 15

20 Glu Pro Ser Pro Ser Leu Gln Gly Cys Ser Glu Ser Pro Val Gly Met
20 25 30

25 Asp

30 (2) INFORMATION FOR SEQ ID NO: 501:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 501:

Met Gln Phe Leu Leu Thr Ala Phe Leu Leu Val Pro Leu Leu Ala Leu
1 5 10 15

40 Cys Asp Val Pro Ile Ser Leu Gly Phe Ser Pro Ser
20 25

45 (2) INFORMATION FOR SEQ ID NO: 502:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 502:

Pro Gly Lys Pro Gln Ala Cys Pro Glu Leu Thr Ser Val Leu Pro
1 5 10 15

55

(2) INFORMATION FOR SEQ ID NO: 503:

60 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 503:

5

Asn Lys Ser Leu Xaa Ser Cys Leu Phe Val Leu His Phe Val Leu His
1 5 10 15

Cys Xaa Phe

10

(2) INFORMATION FOR SEQ ID NO: 504:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 504:

Met Glu Lys Thr His Arg Leu Arg Ile Arg Asn Pro Cys Leu Gln Phe
1 5 10 15

25

Ser Ile Leu Asn Leu Phe Leu Leu Lys Met Ile Val Ser
20 25

30

(2) INFORMATION FOR SEQ ID NO: 505:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 75 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 505:

Met Val Asp Ile Ser Lys Met His Met Ile Leu Tyr Asp Leu Gln Gln
1 5 10 15

40

Asn Leu Ser Ser Ser His Arg Ala Leu Glu Lys Gln Ile Asp Thr Leu
20 25 30

45

Ala Gly Lys Leu Asp Ala Leu Thr Glu Leu Leu Ser Thr Ala Leu Gly
35 40 45

Pro Ser Ser Phe Gln Asn Pro Ala Ser Ser Pro Ser Ser Trp Thr His
50 55 60

50

Glu Glu Glu Pro Gly Tyr Phe Pro Gln Tyr Xaa
65 70 75

55

(2) INFORMATION FOR SEQ ID NO: 506:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 506:

Leu Pro Leu Ala Glu Leu Lys Asn Trp Val
1 5 10

5

(2) INFORMATION FOR SEQ ID NO: 507:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 207 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 507:

15

Met Leu Trp Phe Gly Gly Cys Ser Ala Val Asn Ala Thr Gly His Leu
1 5 10 15

20

Ser Asp Thr Leu Trp Leu Ile Pro Ile Thr Phe Leu Thr Ile Gly Tyr
20 25 30

Gly Asp Val Val Pro Gly Thr Met Trp Gly Lys Ile Val Cys Leu Cys
35 40 45

25

Thr Gly Val Met Gly Val Cys Cys Thr Ala Leu Leu Val Ala Val Val
50 55 60

30

Ala Arg Lys Leu Glu Phe Asn Lys Ala Glu Lys His Val His Asn Phe
65 70 75 80

Met Met Asp Ile Gln Tyr Thr Lys Glu Met Lys Glu Ser Ala Ala Arg
85 90 95

35

Val Leu Gln Glu Ala Trp Met Phe Tyr Lys His Thr Arg Arg Lys Glu
100 105 110

Ser His Ala Ala Arg Arg His Gln Arg Xaa Leu Leu Ala Ala Ile Asn
115 120 125

40

Ala Phe Arg Gln Val Arg Leu Lys His Arg Lys Leu Arg Glu Gln Val
130 135 140

Asn Ser Met Val Asp Ile Ser Lys Met His Met Ile Leu Tyr Asp Leu
145 150 155 160

45

Gln Gln Asn Leu Ser Ser Ser His Arg Ala Leu Glu Lys Gln Ile Asp
165 170 175

50

Thr Leu Ala Gly Lys Leu Asp Ala Leu Thr Glu Leu Leu Ser Thr Ala
180 185 190

Leu Gly Pro Arg Gln Leu Pro Glu Pro Ser Gln Gln Ser Lys Xaa
195 200 205

55

(2) INFORMATION FOR SEQ ID NO: 508:

60

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 508:

5 Met Trp Arg Cys Arg Gly Lys Leu Ser Phe Pro Leu Phe Ala Val Val
1 5 10 15
Ile Val Ser Cys Arg Lys Asp Gly Pro Asp Ala Ala Ala Pro Ala
20 25 30
10 Val Xaa Lys Lys
35

15

(2) INFORMATION FOR SEQ ID NO: 509:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 509:

25 Met Ala Leu Val Ala Leu Phe Thr Gln Leu Met Arg Xaa Leu Gly Arg
1 5 10 15
Cys Pro Gln

30

(2) INFORMATION FOR SEQ ID NO: 510:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 510:

40 Met Thr Phe Pro Phe Glu Lys Glu Asn Ser Cys Phe Gln Cys Leu Leu
1 5 10 15
Phe Asp Ser Trp Arg Glu Gln Thr Arg Thr Asn Ile Gln Pro Gln Arg
20 25 30
45

50

(2) INFORMATION FOR SEQ ID NO: 511:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 511:

60 Met His Leu Leu Asp Phe Phe Arg Asp Leu Val Leu Leu Val Leu Leu
1 5 10 15

Ala Leu Leu Asp Ser Phe Trp Leu Glu Val Gln Lys
20 25

5

(2) INFORMATION FOR SEQ ID NO: 512:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 512:

15

Met Cys Leu Ile His Phe Ile Lys Ile Ile Leu Val Phe Ile Leu Lys
1 5 10 15

Leu Trp Leu Tyr Ser Gln Lys Cys Pro Lys
20 25

20

(2) INFORMATION FOR SEQ ID NO: 513:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 513:

30

Met Ile His Val His Glu Trp Asn Asp Gln Met Leu Met Val Tyr Ile
1 5 10 15

35

Phe Leu Tyr Pro Val Ser Ile Thr Phe Leu Asn Leu Cys Ser Leu Thr
20 25 30

Cys

40

(2) INFORMATION FOR SEQ ID NO: 514:

45

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 514:

50

Leu Asn Glu Ser Tyr Val Ser Arg Ala Gly Gly Trp Phe Ser Met Phe
1 5 10 15

Xaa Leu Ile Phe Phe Leu Leu Ala Leu Gly Ser Xaa Leu Cys Leu Leu
20 25 30

55

Leu Cys Leu Pro Ser Phe Asn Lys Thr Arg Arg Lys Gln Lys Pro
35 40 45

60

(2) INFORMATION FOR SEQ ID NO: 515:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 515:

5 Ser Ser Lys Thr Pro Leu Pro Ser Glu Arg Arg Trp Ile Ser Gly Ser
 10 1 5 10 15
 Ser Leu Met Ala Pro Arg Pro Trp Leu Leu Gly Ile Ala Leu Leu Gly
 20 25 30
 15 Leu Trp Ala Leu Glu Pro Ala Leu Gly His Trp
 35 40

20 (2) INFORMATION FOR SEQ ID NO: 516:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 516:

25 Leu Asn Trp
 30 1

(2) INFORMATION FOR SEQ ID NO: 517:

35 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 174 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 517:

40 Phe Ala Phe Cys Ala Glu Leu Met Ile Gln Asn Trp Thr Leu Gly Ala
 1 5 10 15
 45 Val Asp Ser Gln Met Asp Asp Met Asp Met Asp Leu Asp Lys Glu Phe
 20 25 30
 Leu Gln Asp Leu Lys Glu Leu Lys Val Leu Val Ala Asp Lys Asp Leu
 35 40 45
 50 Leu Asp Leu His Lys Ser Leu Val Cys Thr Ala Leu Arg Gly Lys Leu
 50 55 60
 Gly Val Phe Ser Glu Met Glu Ala Asn Phe Lys Asn Leu Ser Arg Gly
 65 70 75 80
 55 Leu Val Asn Val Ala Ala Lys Leu Thr His Asn Lys Asp Val Arg Asp
 85 90 95
 60 Leu Phe Val Asp Leu Val Glu Lys Phe Val Glu Pro Cys Arg Ser Asp
 100 105 110

His Trp Pro Leu Ser Asp Val Arg Phe Phe Leu Asn Gln Tyr Ser Ala
115 120 125

5 Ser Val His Ser Leu Asp Gly Phe Arg His Gln Ala Ser Gly Thr Ala
130 135 140

Thr Trp Ala Pro Ser Ala Ala Ala Ser Cys Ala Cys Ile Met Thr Glu
145 150 155 160

10 Val Pro Pro Asn Ala Pro Pro Thr Leu Thr Ile Lys Leu Leu
165 170

15

(2) INFORMATION FOR SEQ ID NO: 518:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 518:

20 Met Trp Lys Asn Leu Gly Ser Gly Ser Val Phe Val Thr Trp Phe Ser
1 5 10 15

Leu Val Met Ile Leu Ser Gly Ile Gly Pro Leu Gly Asp Ala Glu Asp
20 25 30

30 Ser Ile Ser Asp Val Ser His Arg Leu Arg Pro
35 40

35

(2) INFORMATION FOR SEQ ID NO: 519:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 519:

40 Phe Gln Phe Pro Leu Leu Thr Ile Ala Leu Gln Phe Leu
1 5 10

45

(2) INFORMATION FOR SEQ ID NO: 520:

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 520:

55

Met His Tyr Val Ile Val Leu Ser Leu Phe Val Val Leu Glu Lys Lys
1 5 10 15

60 Asn Lys Met Gly Ser Asp Gly Cys Leu Arg Lys Asn Gly Ser
20 25 30

(2) INFORMATION FOR SEQ ID NO: 521:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 521:

Met Ser Arg Ser Ile Val Leu Arg Gly Ser Leu Phe Leu Phe Phe Ser
1 5 10 15

15 His Tyr Thr Leu Lys Leu Leu Ser Val Ile Lys Gln Thr Asn Arg Lys
20 25 30

Ile Val Trp Glu Lys Pro Cys Ile Arg Leu Phe Tyr Xaa Val Leu
35 40 45

20

(2) INFORMATION FOR SEQ ID NO: 522:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 522:

Met Pro Leu Pro Val Leu Leu Cys Leu Thr Leu Pro Met Pro Leu Pro
1 5 10 15

35 Ser Ala Thr Ala Arg Gly Gly Asn Arg Thr
20 25

(2) INFORMATION FOR SEQ ID NO: 523:

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 58 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 523:

Ser Ser Ile Pro Val Ser Ile Leu Ile Gly Met Lys Leu Ile Leu Tyr
1 5 10 15

50 Leu Leu Ile Thr Glu Ser Gly Ser His Glu Lys Lys Ser Phe Tyr Pro
20 25 30

Ser Phe Lys Tyr Met Phe Lys Ile Ile Ile Tyr Val Ser Ala Tyr Cys
35 40 45

55

Arg Thr Ala Leu Arg Ala Thr Val Ser His
50 55

60

(2) INFORMATION FOR SEQ ID NO: 524:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 524:

5 Asn Arg Thr Leu Leu Phe Leu Ile Leu Phe Val Leu Phe Gly Leu Gly
10 1 5 10 15
Tyr Gly Phe

15

(2) INFORMATION FOR SEQ ID NO: 525:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 40 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 525:

25 Met Phe Leu Leu Val Leu Ser Val Phe Cys Asp Phe Met Cys Ser Ile
1 5 10 15
Ala Pro Arg Cys His Ala Leu Ser Leu Val Ser Leu Arg Ala Gln His
30 20 25 30
Leu Ser Leu Phe Ile Thr Cys His
35 40

35

(2) INFORMATION FOR SEQ ID NO: 526:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 57 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 526:

45 Met Leu Leu Phe Ile Leu Leu Thr Leu Ser Ser Gly Cys Arg Leu Leu
1 5 10 15
Val Ser Ser Trp Lys Thr Phe Leu Pro His Phe Ser Leu Pro Gly Pro
20 25 30
50 Arg Glu His Pro Glu Gly Ser Arg Thr Trp Phe Phe Arg Tyr Trp Glu
35 40 45
Pro Gly Ala His Cys Leu His Cys Ala
55 50 55

55

(2) INFORMATION FOR SEQ ID NO: 527:

60

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 527:

5

Ala Arg Leu Leu Leu Phe Leu Ser Ser Val His Pro Ser Ile Met Pro
 1 5 10 15

Ser Cys Asn Gln Leu
 10 20

(2) INFORMATION FOR SEQ ID NO: 528:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 528:

Met Ser Leu Thr Ser Ser Leu Thr Phe Leu Ser His Ile Leu Leu Leu
 1 5 10 15

25 Pro Gln Lys Leu Gln Phe Leu Ser Trp Met Glu Arg Gln Gln Arg Cys
 20 25 30

Thr Gly Val Ala Lys Tyr Ala
 30 35

(2) INFORMATION FOR SEQ ID NO: 529:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 128 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 529:

Met Val Leu Arg Leu Ile Gln Leu Ile Phe Leu Ile Phe Phe Ile His
 1 5 10 15

45 Ile Ile Ile Leu Leu Ile Pro Gly Ser Arg Pro Cys Gly Ser Trp Val
 20 25 30

Asn Asp Arg Xaa Leu Gly Leu Arg Asp Val Thr His Leu Ile Tyr Leu
 35 40 45

50 His Trp Val His Gly His Leu Pro Trp Cys His Pro Tyr Ile Gln Val
 50 55 60

Glu Phe Ser Ala Leu Ile Glu Ser Thr Ala Gln Leu Gly Leu Pro Phe
 65 70 75 80

55 Ser Trp Val Arg Val Ile His Pro Phe Leu Val Leu Pro Cys Leu Tyr
 85 90 95

60 Ser Pro Gly Leu Lys Asn Gly Ile Phe Leu Phe Leu Leu Arg Ala Met
 100 105 110

Pro Gly Gly Met Phe Pro Gly Asn Leu Glu Ala Phe Arg Val Pro Val
115 120 125

5

10 (2) INFORMATION FOR SEQ ID NO: 530:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 82 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 530:

Met Gly Ser Ser Val Leu Pro Phe Cys Val Cys Val Thr Ser Pro Ser
1 5 10 15

20

Leu Gly Gly Arg Cys Ile Gln Gly Arg Phe Ala Ser His Ser Lys Phe
20 25 30

25

Trp Gly Phe Gly Xaa Lys Thr Ala Ser Phe Gly Ala Val Gly Glu Thr
35 40 45

Pro Pro Asp Gln Glu Pro Gln Lys Glu Thr Glu Pro Ala Thr Ser Ser
50 55 60

30

His Ala Arg Pro Trp Ala Arg Val Ile Gly Leu Arg Ile Trp Pro Gln
65 70 75 80

Pro Asn

35

(2) INFORMATION FOR SEQ ID NO: 531:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 531:

45

Met Leu Leu Ser Val Ala Ile Phe Ile Leu Leu Thr Leu Val Tyr Ala
1 5 10 15

Tyr Trp Thr Met
20

50

(2) INFORMATION FOR SEQ ID NO: 532:

55

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 75 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 532:

Asn Cys Glu Ile Leu Glu Tyr Cys Tyr Tyr Leu Thr Gln Leu Lys Ile
 1 5 10 15
 5 Ser Met Gly Lys Tyr Leu Ser Ile Pro Thr Val Leu Leu Lys Ile Ile
 20 25 30
 Arg Cys Ser Ile Thr Ala Val Ser Asp Ser Ser Thr Ser Trp Ala Ile
 35 40 45
 10 Lys Ala Gln Leu Lys Ile Glu Asn Lys Asp Leu Asp Asn Lys Thr Ala
 50 55 60
 Lys Gly Gly Gly Gln Glu Ala Leu Thr Cys Thr
 15 65 70 75

(2) INFORMATION FOR SEQ ID NO: 533:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 533:

Met Phe Leu Met Arg Met His Leu Cys Phe Cys Lys Tyr Cys Cys Ser
 1 5 10 15
 30 Phe Ile Val Thr Pro Thr Ser Thr Ser Asn Thr Xaa Ser Tyr Leu Trp
 20 25 30
 Pro Trp Ile Ser Ala Ser Met Ala Gly Arg Gly Ser Xaa Trp Ala Cys
 35 40 45
 35 Thr Leu Asn Ala Val Thr Arg Glu Gly Leu Pro Glu
 50 55 60

40

(2) INFORMATION FOR SEQ ID NO: 534:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 534:

Met Ser Leu Leu Asn Thr His Thr Leu Cys Phe Val Leu Phe Cys Phe
 1 5 10 15
 50 Thr Leu Ser Ile Asn Gln Glu Lys Leu Ala Asn His Leu Ala Phe Arg
 20 25 30
 55 Ile Leu Phe Phe Ile Val Phe
 35

60

(2) INFORMATION FOR SEQ ID NO: 535:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 535:

Met Leu
1

10

- (2) INFORMATION FOR SEQ ID NO: 536:

- 15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 536:

Met Asp Gln Phe Lys Ile Phe Tyr Phe Leu Lys Ala Phe Phe Ala Cys
1 5 10 15

25 Cys Asn Val Gln Asp Pro Ser Pro Phe Met Gly Glu Thr Gly Ser Tyr
20 25 30

Leu Asn Ile Gly
35

30

- (2) INFORMATION FOR SEQ ID NO: 537:

- 35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 537:

40 Met Phe Asp Phe Leu Ser Tyr Phe Lys Asp Leu Leu Ser Cys
1 5 10

- 45 (2) INFORMATION FOR SEQ ID NO: 538:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 538:

Met Gly Phe Gly Phe Val Leu Asn Ile Phe Ser Phe Phe Leu Xaa Pro
1 5 10 15

55

Pro Leu

60

(2) INFORMATION FOR SEQ ID NO: 539:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 539:

Leu Leu Leu Trp Thr Leu Leu Ala Xaa Tyr Xaa
1 5 10

(2) INFORMATION FOR SEQ ID NO: 540:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 108 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 540:

Met Ala Ala Gln Lys Asp Gln Gln Lys Asp Ala Glu Ala Glu Gly Leu
1 5 10 15
Ser Gly Thr Thr Leu Leu Pro Lys Leu Ile Pro Ser Gly Ala Gly Arg
20 25 30
Glu Trp Leu Glu Arg Arg Arg Ala Thr Ile Arg Pro Trp Ser Thr Phe
35 40 45
Val Asp Gln Gln Arg Phe Ser Arg Pro Arg Asn Leu Gly Glu Leu Cys
50 55 60
Gln Arg Leu Val Arg Asn Val Glu Tyr Tyr Gln Ser Asn Tyr Val Phe
65 70 75 80
Val Phe Leu Gly Leu Ile Leu Tyr Cys Val Val Thr Ser Pro Met Leu
85 90 95
Leu Val Ala Leu Ala Val Phe Phe Gly Ala Cys Xaa
100 105

(2) INFORMATION FOR SEQ ID NO: 541:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 106 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 541:

Phe Val Phe Leu Gly Leu Ile Leu Tyr Cys Val Val Thr Ser Pro Met
1 5 10 15
Leu Leu Val Ala Leu Ala Val Phe Phe Gly Ala Cys Tyr Ile Leu Tyr
20 25 30
Leu Arg Thr Leu Glu Ser Lys Leu Val Leu Phe Gly Arg Glu Val Ser
35 40 45

Pro Ala His Gln Tyr Ala Leu Ala Gly Gly Ile Ser Phe Pro Phe Phe
 50 55 60

5 Trp Leu Ala Gly Ala Gly Ser Ala Val Phe Trp Val Leu Gly Ala Thr
 65 70 75 80

Leu Val Val Ile Gly Ser His Ala Ala Phe His Gln Ile Glu Ala Val
 85 90 95

10 Asp Gly Glu Glu Leu Gln Met Glu Pro Val
 100 105

15

(2) INFORMATION FOR SEQ ID NO: 542:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 136 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 542:

20

25 Met Asp Arg Phe Thr Val Ala Gly Val Leu Pro Asp Ile Glu Gln Phe
 1 5 10 15

Phe Asn Ile Gly Asp Ser Ser Ser Gly Leu Ile Gln Thr Val Phe Ile
 20 25 30

30 Ser Ser Tyr Met Val Leu Ala Pro Val Phe Gly Tyr Leu Gly Asp Arg
 35 40 45

Tyr Asn Arg Lys Tyr Leu Met Cys Gly Gly Ile Ala Phe Trp Ser Leu
 50 55 60

35 Val Thr Leu Gly Ser Ser Phe Ile Pro Gly Glu His Phe Trp Leu Leu
 65 70 75 80

40 Leu Leu Thr Arg Gly Leu Val Gly Val Gly Glu Ala Ser Tyr Ser Thr
 85 90 95

Ile Ala Pro Thr Leu Ile Ala Asp Leu Phe Val Ala Asp Gln Arg Thr
 100 105 110

45 Gly Cys Ser Ala Ser Ser Thr Leu Pro Phe Arg Trp Ala Val Val Trp
 115 120 125

Ala Thr Leu Gln Ala Pro Lys Xaa
 130 135

50

(2) INFORMATION FOR SEQ ID NO: 543:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 424 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 543:

55

60

Met Ala Gly Asp Trp His Trp Ala Leu Arg Val Thr Pro Gly Leu Gly
 1 5 10 15
 5 Val Val Ala Val Leu Leu Leu Phe Leu Val Val Arg Glu Pro Pro Arg
 20 25 30
 Gly Ala Val Glu Arg His Ser Asp Leu Pro Pro Leu Asn Pro Thr Ser
 35 40 45
 10 Trp Trp Ala Asp Leu Arg Ala Leu Ala Arg Asn Pro Ser Phe Val Leu
 50 55 60
 Ser Ser Leu Gly Phe Thr Ala Val Ala Phe Val Thr Gly Ser Leu Ala
 65 70 75 80
 15 Leu Trp Ala Pro Ala Phe Leu Leu Arg Ser Arg Val Val Leu Gly Glu
 85 90 95
 Thr Pro Pro Cys Leu Pro Gly Asp Ser Cys Ser Ser Ser Asp Ser Leu
 100 105 110
 20 Ile Phe Gly Leu Ile Thr Cys Leu Thr Gly Val Leu Gly Val Gly Leu
 115 120 125
 25 Gly Val Glu Ile Ser Arg Arg Xaa Arg His Ser Asn Pro Arg Ala Asp
 130 135 140
 Pro Leu Val Cys Ala Thr Gly Leu Leu Gly Ser Ala Pro Phe Leu Phe
 145 150 155 160
 30 Leu Ser Leu Ala Cys Ala Arg Gly Ser Ile Val Ala Thr Tyr Ile Phe
 165 170 175
 35 Ile Phe Ile Gly Glu Thr Leu Leu Ser Met Asn Trp Ala Ile Val Ala
 180 185 190
 Asp Ile Leu Leu Tyr Val Val Ile Pro Thr Arg Arg Ser Thr Ala Glu
 195 200 205
 40 Ala Phe Gln Ile Val Leu Ser His Leu Leu Gly Asp Ala Gly Ser Pro
 210 215 220
 Tyr Leu Ile Gly Leu Ile Ser Asp Arg Leu Arg Arg Asn Trp Pro Pro
 225 230 235 240
 45 Ser Phe Leu Ser Glu Phe Arg Ala Leu Gln Phe Ser Leu Met Leu Cys
 245 250 255
 50 Ala Phe Val Gly Ala Leu Gly Gly Ala Leu Ser Trp Ala Pro Xaa Ser
 260 265 270
 Ser Leu Arg Pro Thr Ala Gly Gly His Ser Cys Thr Cys Arg Ala Cys
 275 280 285
 55 Cys Thr Lys Gln Gly Pro Gln Thr Thr Gly Leu Trp Cys Pro Ser Gly
 290 295 300
 60 Ala Ala Pro Pro Ala Cys Pro Trp Pro Val Cys Ser Ser Glu Arg Leu
 305 310 315 320

Pro Leu Thr Tyr Leu His Ile Cys His Ser Xaa Pro Trp Ala His Pro
 325 330 335

5 Thr Lys Gly Leu Gly Leu Thr Pro Trp Pro Gly Pro Ala Ser Arg Gly
 340 345 350

Thr Leu Gly Arg Val Pro Ala Pro Arg His Tyr Xaa Gly Ser Ser Gly
 355 360 365

10 Glu Glu Val Gly Val Gln Glu Gly Asp Pro Ser Pro Gln Gly Xaa Pro
 370 375 380

Gln Gly Leu Gly Ala Ile Cys Asn Gly Ile Lys Phe Val Ala Arg Pro
 385 390 395 400

15 Gln Val Pro Ala Leu Val Phe Leu Trp Val Ala Ser Asp Leu Ala Pro
 405 410 415

20 Arg Leu His Pro Arg Ala Pro Glu
 420

25 (2) INFORMATION FOR SEQ ID NO: 544:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 39 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 544:

Met Phe Arg Phe Val Ile Cys Leu Phe Leu Trp Leu Val Leu Cys Arg
 1 5 10 15

35 Asp Ser Thr Ser Ala Ser Arg Ile Ala Leu Tyr Tyr Arg Ile Val Phe
 20 25 30

Leu Ile His Gln Cys Ser Ser
 35

40

(2) INFORMATION FOR SEQ ID NO: 545:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 58 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 545:

Met Leu Pro Trp Xaa Ala Gln Leu Leu Asp Arg Thr Ile Gly Pro Leu
 1 5 10 15

55 Tyr Leu Leu Phe Val Gln Phe Ser Pro Ala Phe Ser Arg Thr Ser Pro
 20 25 30

Trp Arg Ser Pro Lys Asn Phe Arg Arg Leu Tyr Pro Pro Cys Thr Thr
 35 40 45

60 Ser Gly Cys Ala Ala Arg Trp Leu Phe Ser

50

55

5 (2) INFORMATION FOR SEQ ID NO: 546:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 amino acids

(B) TYPE: amino acid

10 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 546:

Met Gly Leu Ser Val Leu Leu Pro Leu Cys Leu Leu Gly Pro Gly Arg
1 5 10 15
Phe Thr Ser Gly Gln Lys Pro Leu Asp Thr Pro Gly Leu Gly Val Pro
20 25 30

20 Phe

25 (2) INFORMATION FOR SEQ ID NO: 547:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 367 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 547:

Met Ala Lys Pro Gln Val Val Val Ala Pro Val Leu Met Ser Lys Leu
1 5 10 15
Ser Val Asn Ala Pro Glu Phe Tyr Pro Ser Gly Tyr Ser Ser Ser Tyr
20 25 30
Thr Glu Ser Tyr Glu Asp Gly Cys Glu Asp Tyr Pro Thr Leu Ser Glu
35 40 45
Tyr Val Gln Asp Phe Leu Asn His Leu Thr Glu Gln Pro Gly Ser Phe
50 55 60
Glu Thr Glu Ile Glu Gln Phe Ala Glu Thr Leu Asn Gly Cys Val Thr
45 65 70 75 80
Thr Asp Asp Ala Leu Gln Glu Leu Val Glu Leu Ile Tyr Gln Gln Ala
85 90 95
Thr Ser Ile Pro Asn Phe Ser Tyr Met Gly Ala Arg Leu Cys Asn Tyr
50 100 105 110
Leu Ser His His Leu Thr Ile Ser Pro Gln Ser Gly Asn Phe Arg Gln
115 120 125
Leu Leu Leu Gln Arg Cys Arg Thr Glu Tyr Glu Val Lys Asp Gln Ala
55 130 135 140
Ala Lys Gly Asp Glu Val Thr Arg Lys Arg Phe His Ala Phe Val Leu
60 145 150 155 160

Phe Leu Gly Glu Leu Tyr Leu Asn Leu Glu Ile Lys Gly Thr Asn Gly
 165 170 175
 5 Gln Val Thr Arg Ala Asp Ile Leu Gln Val Gly Leu Arg Glu Leu Leu
 180 185 190
 Asn Ala Leu Phe Ser Asn Pro Met Asp Asp Asn Leu Ile Cys Ala Val
 195 200 205
 10 Lys Leu Leu Lys Leu Thr Gly Ser Val Leu Glu Asp Ala Trp Lys Glu
 210 215 220
 Lys Gly Lys Met Asp Met Glu Glu Ile Ile Gln Arg Ile Glu Asn Val
 15 225 230 235 240
 Val Leu Asp Ala Asn Cys Ser Arg Asp Val Lys Gln Met Leu Leu Lys
 245 250 255
 20 Leu Val Glu Leu Arg Ser Ser Asn Trp Gly Arg Val His Ala Thr Ser
 260 265 270
 Thr Tyr Arg Glu Ala Thr Pro Glu Asn Asp Pro Asn Tyr Phe Met Asn
 275 280 285
 25 Glu Pro Thr Phe Tyr Thr Ser Asp Gly Val Pro Phe Thr Ala Ala Asp
 290 295 300
 Pro Asp Tyr Gln Glu Lys Tyr Gln Glu Leu Leu Glu Arg Glu Asp Phe
 30 305 310 315 320
 Phe Pro Asp Tyr Glu Glu Asn Gly Thr Asp Leu Ser Gly Ala Gly Asp
 325 330 335
 35 Pro Tyr Leu Asp Asp Ile Asp Asp Glu Met Asp Pro Glu Ile Glu Glu
 340 345 350
 Ala Tyr Glu Lys Phe Cys Leu Glu Ser Glu Arg Lys Arg Lys Gln
 355 360 365
 40

(2) INFORMATION FOR SEQ ID NO: 548:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 77 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 548:
 50 Met Leu Arg Leu Asp Ile Ile Asn Ser Leu Val Thr Thr Val Phe Met
 1 5 10 15
 Leu Ile Val Ser Val Leu Ala Leu Ile Pro Glu Thr Thr Thr Leu Thr
 55 20 25 30
 Val Gly Gly Gly Val Phe Ala Leu Val Thr Ala Val Cys Cys Leu Ala
 35 40 45
 60 Asp Gly Ala Leu Ile Tyr Arg Lys Leu Leu Phe Asn Pro Ser Gly Pro

Tyr Gln Lys Lys Pro Val His Glu Lys Lys Glu Val Leu
65 70 75

5

(2) INFORMATION FOR SEQ ID NO: 549:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 549:

15

Met Leu Lys Gln Val Met Phe Val Phe Ser Gly Met Gly Pro Arg Ser
1 5 10 15

20

His Cys Trp Gly Leu Pro Leu Ala Cys Gly Thr Phe Val Gln Gly His
20 25 30

Gln Ala Asp Ser Ser His Leu Leu Pro Leu Lys His Gln Gly Ala
35 40 45

25

(2) INFORMATION FOR SEQ ID NO: 550:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 168 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 550:

35

Met Leu Leu Ser Leu Ala Ala Phe Ser Val Ile Ser Val Val Ser Tyr
1 5 10 15

Leu Ile Leu Ala Leu Leu Ser Val Thr Ile Ser Phe Arg Ile Tyr Lys
20 25 30

40

Ser Val Ile Gln Ala Val Gln Lys Ser Glu Glu Gly His Pro Phe Lys
35 40 45

45

Ala Tyr Leu Asp Val Asp Ile Thr Leu Ser Ser Glu Ala Phe His Asn
50 55 60

Tyr Met Asn Ala Ala Met Val His Ile Asn Arg Ala Leu Lys Leu Ile
65 70 75 80

50

Ile Arg Leu Phe Leu Val Glu Asp Leu Val Asp Ser Leu Lys Leu Ala
85 90 95

Val Phe Met Trp Leu Met Thr Tyr Val Gly Ala Val Phe Asn Gly Ile
100 105 110

55

Thr Leu Leu Ile Leu Ala Glu Leu Leu Ile Phe Ser Val Pro Ile Val
115 120 125

60

Tyr Glu Lys Tyr Lys Thr Gln Ile Asp His Tyr Val Gly Ile Ala Arg
130 135 140

Asp Gln Thr Lys Ser Ile Val Glu Lys Ile Gln Ala Lys Leu Pro Gly
145 150 155 160

5 Ile Ala Lys Lys Lys Ala Glu Xaa
165

10 (2) INFORMATION FOR SEQ ID NO: 551:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 124 amino acids

(B) TYPE: amino acid

15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 551:

Ser Val Pro Phe His Leu Leu Val Val Leu Arg Ser Arg Ala Val Arg
1 5 10 15
20 Ala Arg Arg Arg Arg Glu Pro Arg Ser Leu Pro Arg Pro Gly Asp Glu
20 25 30
25 Glu Leu Gln Leu Leu Leu Cys Gly Ala Arg Ser Asp Phe Leu Glu Arg
35 40 45
Cys Glu Glu Asp Trp Val Cys Leu Trp His His Ala Asp His Ala Ala
50 55 60
30 Phe Pro Gly Ser Phe Gln Cys His Gln Cys Gly Phe Leu Pro His Pro
65 70 75 80
Gly Ser Ser Leu Cys His His Gln Leu Gln Asp Leu Gln Val Arg His
85 90 95
35 Pro Ser Cys Thr Glu Val Arg Arg Arg Pro Ser Ile Gln Ser Leu Pro
100 105 110
40 Gly Arg Arg His Tyr Ser Val Leu Arg Ser Phe Pro
115 120

45 (2) INFORMATION FOR SEQ ID NO: 552:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 177 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 552:

Met Val His Leu Leu Val Leu Ser Gly Ala Trp Gly Met Gln Met Trp
1 5 10 15
55 Val Thr Phe Val Ser Gly Phe Leu Leu Phe Arg Ser Leu Pro Arg His
20 25 30
Thr Phe Gly Leu Val Gln Ser Lys Leu Phe Pro Phe Tyr Phe His Ile
35 40 45
60

Ser Met Gly Cys Ala Phe Ile Asn Leu Cys Ile Leu Ala Ser Gln His
 50 55 60
 5 Ala Trp Ala Gln Leu Thr Phe Trp Glu Ala Ser Gln Leu Tyr Leu Leu
 65 70 75 80
 Phe Leu Ser Leu Thr Leu Ala Thr Val Asn Ala Arg Trp Leu Glu Pro
 85 90 95
 10 Arg Thr Thr Ala Ala Met Trp Ala Leu Gln Thr Val Glu Lys Glu Arg
 100 105 110
 Gly Leu Gly Gly Glu Val Pro Gly Ser His Gln Gly Pro Asp Pro Tyr
 115 120 125
 15 Arg Gln Leu Arg Glu Lys Asp Pro Lys Tyr Ser Ala Leu Arg Gln Asn
 130 135 140
 Phe Phe Arg Tyr His Gly Leu Ser Ser Leu Cys Asn Leu Gly Cys Val
 145 150 155 160
 Leu Ser Asn Gly Leu Cys Leu Ala Gly Leu Ala Leu Glu Ile Arg Ser
 165 170 175
 25 Leu

30 (2) INFORMATION FOR SEQ ID NO: 553:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 72 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 553:

Met Ala Phe Ile Leu Leu Phe Tyr Cys Leu Met Thr Phe Leu Ser Leu
 1 5 10 15
 40 Glu Gln Asn Ser Ala Thr Val Glu Pro Ser Ser His Glu Ile Leu His
 20 25 30
 45 Leu Leu Gln Asn Cys Phe Glu Leu Leu Arg Thr Ser Thr Ser Gln Cys
 35 40 45
 Thr Glu Gly Ile Pro Cys Ala Lys Ile Pro Glu Trp Val Thr His Leu
 50 55 60
 50 Thr Trp Gln Thr Leu Lys Asn Ser
 65 70

55 (2) INFORMATION FOR SEQ ID NO: 554:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 45 amino acids

(B) TYPE: amino acid

60 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 554:

5 Val Leu Arg Ile Ile Cys Leu Trp Pro Cys Gly Thr Thr Leu Pro Leu
1 5 10 15
Val Glu Lys Ala His Asp Ser His Ser Ala Asp Pro Val Cys Pro Gly
20 25 30
10 Leu Thr Ala His Leu Pro Val Leu Leu Tyr Val Gln Leu
35 40 45

(2) INFORMATION FOR SEQ ID NO: 555:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 251 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 555:

Met Lys His Ala Asp Pro Arg Ile Gln Gly Tyr Pro Leu Met Gly Ser
1 5 10 15
25 Pro Leu Leu Met Thr Ser Ile Leu Leu Thr Tyr Val Tyr Phe Val Leu
20 25 30
Ser Leu Gly Pro Arg Ile Met Ala Asn Arg Lys Pro Phe Gln Leu Arg
35 40 45
30 Gly Phe Met Ile Val Tyr Asn Phe Ser Leu Val Ala Leu Ser Leu Tyr
50 55 60
Ile Val Tyr Glu Phe Leu Met Ser Gly Trp Leu Ser Thr Tyr Thr Trp
35 65 70 75 80
Arg Cys Asp Pro Gln Asp Cys Thr Leu Gly Gln Cys Pro Ser Val Pro
85 90 95
40 Ser Pro Xaa Thr Pro Val Thr Lys Ala Tyr Val Val Arg Thr Glu Gln
100 105 110
Gly Thr Gly Pro Pro Leu Pro Thr Ala Ala Leu Gln Gly Pro Arg Leu
115 120 125
45 Trp Phe Leu Thr His Phe Pro Arg Ala Ala Pro Gly Met Trp Pro His
130 135 140
Cys Cys Leu Pro Leu Gln Ser Trp Gly Leu Lys Gly Leu Tyr Ser Tyr
50 145 150 155 160
Phe Pro Leu Pro Ala Leu Lys Leu Gly Arg Gly Ala Leu Arg Ala Gly
165 170 175
55 Pro Thr Lys Gly Leu Val Ala Phe Phe Leu Thr Gln Lys Arg Ser Ala
180 185 190
Ile Met Ser Leu Trp Thr Gln Ser His Ser Ser Thr Pro His Thr Glu
195 200 205
60

Ala Val Ala Ser Gly Pro Lys Val Arg Val Gly Gly Gly Leu Gly Ile
 210 215 220

5 Gln Pro Val Glu Ala Ala Tyr Ser Thr Cys Val Leu Ile Lys Ser Asp
 225 230 235 240

Arg Gly Asn His Glu Lys Lys Lys Lys Lys Lys
 245 250

10

(2) INFORMATION FOR SEQ ID NO: 556:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 19 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 556:

20 Gly Leu Ala Gly Leu Cys Gly Gln Leu Ser Ser Pro Ala Leu Cys Val
 1 5 10 15

Asn Arg Leu

25

(2) INFORMATION FOR SEQ ID NO: 557:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 217 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 557:

35 Met Ile Thr Glu Lys Trp Gly Leu Asn Met Glu Tyr Cys Arg Gly Gln
 1 5 10 15

40 Ala Tyr Ile Xaa Ser Ser Gly Phe Ser Ser Lys Met Lys Val Val Ala
 20 25 30

Ser Arg Leu Leu Glu Lys Tyr Pro Gln Ala Ile Tyr Thr Leu Cys Ser
 35 40 45

45 Ser Cys Ala Leu Asn Met Trp Leu Ala Lys Ser Val Pro Val Met Gly
 50 55 60

Val Ser Val Ala Leu Gly Thr Ile Glu Glu Val Cys Ser Phe Phe His
 65 70 75 80

50

Arg Ser Pro Gln Leu Leu Leu Glu Leu Asp Asn Val Ile Ser Val Leu
 85 90 95

55 Phe Gln Asn Ser Lys Glu Arg Gly Lys Glu Leu Lys Glu Ile Cys His
 100 105 110

Ser Gln Trp Thr Gly Arg His Asp Ala Phe Glu Ile Leu Val Glu Leu
 115 120 125

60

Leu Gln Ala Leu Val Leu Cys Leu Asp Gly Ile Asn Ser Asp Thr Asn

130 135 140

Ile Arg Trp Asn Asn Tyr Ile Ala Gly Arg Ala Phe Val Leu Cys Ser
 145 150 155 160

5 Ala Val Ser Asp Phe Asp Phe Ile Val Thr Ile Val Val Leu Lys Asn
 165 170 175

10 Val Leu Ser Phe Thr Arg Ala Phe Gly Lys Asn Leu Gln Gly Gln Thr
 180 185 190

Ser Asp Val Phe Phe Ala Ala Gly Ser Leu Thr Ala Val Leu His Ser
 195 200 205

15 Leu Asn Glu Val Ile Gly Lys Tyr Xaa
 210 215

20 (2) INFORMATION FOR SEQ ID NO: 558:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 82 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 558:

Leu Leu Lys Val Leu Cys Ile Leu Pro Val Met Lys Val Glu Asn Glu
 1 5 10 15

30 Arg Tyr Glu Asn Gly Arg Lys Arg Leu Lys Ala Tyr Leu Arg Asn Thr
 20 25 30

35 Leu Thr Asp Gln Arg Ser Ser Asn Leu Ala Leu Leu Asn Ile Asn Phe
 35 40 45

Asp Ile Lys His Asp Leu Asp Leu Met Val Asp Thr Tyr Ile Lys Leu
 50 55 60

40 Tyr Thr Ser Lys Ser Glu Leu Pro Thr Asp Asn Ser Glu Thr Val Glu
 65 70 75 80

Asn Thr

45

(2) INFORMATION FOR SEQ ID NO: 559:

50 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 95 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 559:

Met Val Leu Ile Leu Leu Asn Leu Leu Leu Gly Gln Phe Ser Cys Met
 1 5 10 15

60 Ser Pro Ala Ser His His Cys His Pro Leu Pro Thr Glu Met Pro Cys
 20 25 30

Ser Ser Asp Trp Gly Phe Asp Ser His Thr Val Tyr Pro Ser Cys Val
 35 40 45

5 Asp Ala Leu Leu Pro Lys Pro Ser Ala Asn Ser Phe Pro Asn Gly Ser
 50 55 60

Cys His Cys Gln Gly Leu Tyr Asn Gln Gln Gln Gln Asn Leu His Ala
 65 70 75 80

10 Ala Glu Gly Pro Ala Ser Leu Arg Cys Asn Lys Tyr Val Ser Thr
 85 90 95

15

(2) INFORMATION FOR SEQ ID NO: 560:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 54 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 560:

20 Met Ile Pro Ala Tyr Ser Lys Asn Arg Ala Tyr Ala Ile Phe Phe Ile
 1 5 10 15

Val Phe Thr Val Ile Gly Asp Ala Pro Gly Ala Val Leu Ser Cys Ala
 20 25 30

30 Gly His Pro Cys Val Gly Phe Ala Ala Val Leu Val Ala Pro Leu Thr
 35 40 45

Val Ala Val Ser Ser Xaa
 50

35

(2) INFORMATION FOR SEQ ID NO: 561:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 108 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 561:

40

45 Met Glu Val Pro Pro Pro Ala Pro Arg Ser Phe Leu Cys Arg Ala Leu
 1 5 10 15

50 Cys Leu Phe Pro Arg Val Phe Ala Ala Glu Ala Val Thr Ala Asp Ser
 20 25 30

Glu Val Leu Glu Glu Arg Gln Lys Arg Leu Pro Tyr Val Pro Glu Pro
 35 40 45

55 Tyr Tyr Pro Glu Ser Gly Trp Asp Arg Leu Arg Glu Leu Phe Gly Lys
 50 55 60

Asp Thr Val Asn Thr Ser Leu Asn Val Tyr Arg Asn Lys Asp Ala Leu
 65 70 75 80

60

Ser His Phe Val Ile Ala Gly Ala Val Thr Gly Ser Leu Phe Arg Ile
85 90 95

5 Asn Val Gly Leu Arg Gly Trp Trp Leu Val Ala Xaa
100 105

10 (2) INFORMATION FOR SEQ ID NO: 562:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 50 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 562:

Met Asn Trp Gly Leu Ser Ile Trp Leu His Tyr Tyr Glu Lys Lys Lys
1 5 10 15

20 Glu Gln Val Phe Leu Val Ile Leu Ala His Val Val Arg Arg Cys Ala
20 25 30

Ser Asp Gly Ile Leu Gln Phe Glu Ser Ser Leu Leu Lys Met Arg Arg
25 35 40 45

Ala Pro
50

30 (2) INFORMATION FOR SEQ ID NO: 563:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 253 amino acids

35 (B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 563:

40 Met Val Lys Val Cys Asn Asp Ser Asp Arg Trp Ser Leu Ile Ser Leu
1 5 10 15

Ser Asn Asn Ser Gly Lys Asn Val Glu Leu Lys Phe Val Asp Ser Leu
20 25 30

45 Arg Arg Gln Phe Glu Phe Ser Val Asp Ser Phe Gln Ile Lys Leu Asp
35 40 45

Ser Leu Leu Leu Phe Tyr Glu Cys Ser Glu Asn Pro Met Thr Glu Thr
50 55 60

50 Phe His Pro Thr Ile Ile Gly Glu Ser Val Tyr Gly Asp Phe Gln Glu
65 70 75 80

55 Ala Phe Asp His Leu Cys Asn Lys Ile Ile Ala Thr Arg Asn Pro Glu
85 90 95

Glu Ile Arg Gly Gly Gly Leu Leu Lys Tyr Cys Asn Leu Leu Val Arg
100 105 110

60 Gly Phe Arg Pro Ala Ser Asp Glu Ile Lys Thr Leu Gln Arg Tyr Met

685

115 120 125
 Cys Ser Arg Phe Phe Ile Asp Phe Ser Asp Ile Gly Glu Gln Gln Arg
 130 135 140
 5 Lys Leu Glu Ser Tyr Leu Gln Asn His Phe Val Gly Leu Glu Asp Arg
 145 150 155 160
 10 Lys Tyr Glu Tyr Leu Met Thr Leu His Gly Val Val Asn Glu Ser Thr
 165 170 175
 Val Cys Leu Met Gly His Glu Arg Arg Gln Thr Leu Asn Leu Ile Thr
 180 185 190
 15 Met Leu Ala Ile Arg Val Leu Ala Asp Gln Asn Val Ile Pro Asn Val
 195 200 205
 Ala Asn Val Thr Cys Tyr Tyr Gln Pro Ala Pro Tyr Val Ala Asp Ala
 210 215 220
 20 Asn Phe Ser Asn Tyr Tyr Ile Ala Gln Val Gln Pro Val Phe Thr Cys
 225 230 235 240
 25 Gln Gln Gln Thr Tyr Ser Thr Trp Leu Pro Cys Asn Xaa
 245 250

30 (2) INFORMATION FOR SEQ ID NO: 564:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 564:

Met Ser Phe Leu Met Trp Leu Met Ser Leu Ala Ile Thr Ser Gln Pro
 1 5 10 15

40 Pro Met

45 (2) INFORMATION FOR SEQ ID NO: 565:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 80 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 565:

Met Ala Pro Lys Gly Lys Val Gly Thr Arg Gly Lys Lys Gln Ile Phe
 1 5 10 15

55 Glu Glu Asn Arg Glu Thr Leu Lys Phe Tyr Leu Arg Ile Ile Leu Gly
 20 25 30

60 Ala Asn Ala Ile Tyr Cys Leu Val Thr Leu Val Phe Phe Tyr Ser Ser
 35 40 45

Ala Ser Phe Trp Ala Trp Leu Ala Leu Gly Phe Ser Leu Ala Val Tyr
 50 55 60

5 Gly Ala Ser Tyr His Ser Met Ser Ser Met Ala Arg Ala Ala Phe Phe
 65 70 75 80

10

(2) INFORMATION FOR SEQ ID NO: 566:

15 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 73 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 566:

20

His Leu Lys Asp Val Ile Leu Leu Thr Ala Ile Val Gln Val Leu Ser
 1 5 10 15

25 Cys Phe Ser Leu Tyr Val Trp Ser Phe Trp Leu Leu Ala Pro Gly Arg
 20 25 30

Ala Leu Tyr Leu Leu Trp Val Asn Val Leu Gly Pro Trp Phe Thr Ala
 35 40 45

30 Asp Ser Gly Thr Pro Ala Pro Glu His Asn Glu Lys Arg Gln Arg Arg
 50 55 60

Gln Glu Arg Arg Gln Met Lys Arg Leu
 65 70

35

(2) INFORMATION FOR SEQ ID NO: 567:

40 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 263 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 567:

45

Met Asp Cys Pro Ala Leu Pro Pro Gly Trp Lys Lys Glu Glu Val Ile
 1 5 10 15

50 Arg Lys Ser Gly Leu Ser Ala Gly Lys Ser Asp Val Tyr Tyr Phe Ser
 20 25 30

Pro Ser Gly Lys Lys Phe Arg Ser Lys Pro Gln Leu Ala Arg Tyr Leu
 35 40 45

55 Gly Asn Thr Val Asp Leu Ser Ser Phe Asp Phe Arg Thr Gly Lys Met
 50 55 60

Met Pro Ser Lys Leu Gln Lys Asn Lys Gln Arg Leu Arg Asn Asp Pro
 65 70 75 80

60

[illegible]

(2) INFORMATION FOR SEQ ID NO: 569:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 34 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 569:

Met Pro Val Thr Ser Lys Arg Thr Leu Phe Phe Pro Asp Pro Cys Ser
1 5 10 15

15 Tyr Asp Thr Pro Pro Pro Asp Cys His Cys His Ser Phe Arg Ala Glu
20 25 30

Leu Leu

20

(2) INFORMATION FOR SEQ ID NO: 570:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 104 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 570:

Met Asn Ser Arg Gly Met Trp Leu Thr Tyr Ala Leu Gly Val Gly Leu
1 5 10 15

35 Leu His Ile Val Leu Leu Ser Ile Pro Phe Phe Ser Val Pro Val Ala
20 25 30

Trp Thr Leu Thr Asn Ile Ile His Asn Leu Gly Met Tyr Val Phe Leu
35 40 45

40 His Ala Val Lys Gly Thr Pro Phe Glu Thr Pro Asp Gln Gly Lys Ser
50 55 60

Lys Ala Pro Asn Ser Leu Gly Thr Thr Gly Leu Trp Ser Thr Val Tyr
65 70 75 80

45

Ile Phe Thr Glu Val Phe His Asn Phe Ser Asn Asn Ser Ile Phe Ser
85 90 95

Gly Lys Phe Leu Tyr Glu Val Xaa
100

50

(2) INFORMATION FOR SEQ ID NO: 571:

55

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 132 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 571:

Met Trp Leu Thr Tyr Ala Leu Gly Val Gly Leu Leu His Ile Val Leu
 1 5 10 15
 5 Leu Ser Ile Pro Phe Phe Ser Val Pro Val Ala Trp Thr Leu Thr Asn
 20 25 30
 Ile Ile His Asn Leu Gly Met Tyr Val Phe Leu His Ala Val Lys Gly
 35 40 45
 10 Thr Pro Phe Glu Thr Pro Asp Gln Gly Lys Ala Arg Leu Leu Thr His
 50 55 60
 Trp Glu Gln Leu Asp Tyr Gly Val Gln Phe Thr Ser Ser Arg Lys Phe
 15 65 70 75 80
 Phe Thr Ile Ser Pro Ile Ile Leu Tyr Phe Leu Ala Ser Phe Tyr Thr
 85 90 95
 20 Lys Tyr Asp Pro Thr His Phe Ile Leu Asn Thr Ala Ser Leu Leu Ser
 100 105 110
 Val Leu Ile Pro Lys Met Pro Gln Leu His Gly Val Arg Ile Phe Gly
 115 120 125
 25 Ile Asn Lys Tyr
 130

30

(2) INFORMATION FOR SEQ ID NO: 572:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 572:

40 Met Asn Lys Trp Ile Cys Glu Met His Cys Tyr Leu Val Leu Leu Ser
 1 5 10 15
 Val Cys Ser Pro Ser Ala Leu Arg Arg Val Arg His Thr Leu Ser Arg
 20 25 30

45

50

(2) INFORMATION FOR SEQ ID NO: 573:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 28 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 573:

60 Met Pro Val Leu Ser Leu Leu Cys Thr Leu Ile Val Ser Phe Gln Ser
 1 5 10 15

Ala Asp Ser Cys Glu Val Phe Leu Asn Cys Ser Leu
20 25

5

(2) INFORMATION FOR SEQ ID NO: 574:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 40 amino acids

10

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 574:

15 Met Lys Val Ser Thr Met Leu Trp Phe Leu Cys Trp Glu Gln Ser His
1 5 10 15

Phe Leu Arg Glu Trp Glu Asp Leu Ser Thr Phe Leu Ile Leu Ile Gln
20 25 30

20 Met Glu Cys Gln Tyr Gly Asn Ser
35 40

25

(2) INFORMATION FOR SEQ ID NO: 575:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 amino acids

30

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 575:

35 Met Gly Leu Pro Leu Met Ala Leu Met Trp Ser Thr Leu Pro Ala Ser
1 5 10 15

Ala Gly Val Asn Phe Ile Leu Ala Leu Pro Leu Leu Xaa Leu
20 25 30

40

(2) INFORMATION FOR SEQ ID NO: 576:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 amino acids

45

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 576:

50 Met Lys Arg Gly Cys Leu Gly Leu Leu Phe Phe Ser Cys Cys Ser Ser
1 5 10 15

Ala Pro Thr Met Leu Leu Cys Asp Tyr Leu Asn Trp Phe
20 25

55

(2) INFORMATION FOR SEQ ID NO: 577:

(i) SEQUENCE CHARACTERISTICS:

60

(A) LENGTH: 92 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 577:

5 Met Lys Leu Leu Leu Gly Ile Ala Leu Leu Ala Tyr Val Ala Ser Val
 1 5 10 15

Trp Gly Asn Phe Val Asn Met Arg Ser Ile Gln Glu Asn Gly Glu Leu
 20 25 30

10 Lys Ile Glu Ser Lys Ile Glu Glu Met Val Glu Pro Leu Arg Glu Lys
 35 40 45

Ile Arg Asp Leu Glu Lys Ser Phe Thr Gln Lys Tyr Pro Pro Val Lys
 15 50 55 60

Phe Leu Ser Glu Lys Asp Arg Lys Arg Ile Leu Xaa Asn Arg Arg Arg
 65 70 75 80

20 Xaa Val Arg Gly Leu Pro Ser Xaa Leu Thr Asn Ser
 85 90

25 (2) INFORMATION FOR SEQ ID NO: 578:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 42 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 578:

Met Lys Phe Ser Leu Val Leu Leu Ile Lys Ile Ile Ser Phe Glu Arg
 1 5 10 15

35 Leu Leu Ile Phe Leu Phe Pro Leu Ser Phe Leu Pro Asn Ile Trp Arg
 20 25 30

Arg Val Met Val Asn Leu Asn Ile Leu Phe
 40 35 40

45 (2) INFORMATION FOR SEQ ID NO: 579:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 70 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 579:

Leu Ala Gln Glu Cys Pro Pro His Ile Pro Ser Ser Phe Phe Leu Val
 1 5 10 15

55 Lys Leu Leu Phe Ile Pro Trp Leu Ala Ser Leu Leu Pro Pro Leu Ser
 20 25 30

Thr Phe Thr Ser Asp Phe Tyr Phe Met Glu Phe Gly Ile Glu Val Lys
 35 40 45

60

Leu Gln Gln Cys Arg Gln His Gln Val Leu Gln Glu Lys Asn Thr Lys
 50 55 60

5 Lys Phe Asn Lys Lys Lys
 65 70

(2) INFORMATION FOR SEQ ID NO: 580:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 110 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 580:

Met Leu Arg Leu Leu Leu Val Ala Phe Ala Leu Val Val Val Leu
 1 5 10 15

20

Phe His Val Leu Leu Ala Pro Ile Thr Ala Leu Phe His Thr His Phe
 20 25 30

Tyr Asp Arg Leu Gln Asp Ala Gly Ser Arg Trp Pro Glu Leu Tyr Leu
 35 40 45

25

Tyr Ser Arg Ala Asp Glu Val Val Leu Ala Arg Asp Ile Glu Arg Met
 50 55 60

30

Val Glu Ala Arg Leu Ala Arg Arg Val Leu Ala Arg Ser Val Asp Phe
 65 70 75 80

Val Ser Ser Ala His Val Ser His Leu Arg Asp Tyr Pro Thr Tyr Tyr
 85 90 95

35

Thr Ser Leu Cys Val Asp Phe Met Arg Asn Cys Val Arg Cys
 100 105 110

40

(2) INFORMATION FOR SEQ ID NO: 581:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 581:

Met Phe Lys Leu Glu Glu Cys Gly Lys Thr Thr Phe Leu Leu Ser Met
 1 5 10 15

50

Ala Leu Tyr Phe Trp Trp Ile Val Gln Thr Thr Lys Gly Cys
 20 25 30

55

(2) INFORMATION FOR SEQ ID NO: 582:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 71 amino acids

(B) TYPE: amino acid

60

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 582:

5 Met Glu Ser Asp Ala Leu Leu Leu Thr Ile Phe Trp Ile Ile Ala Arg
1 5 10 15
Ser Ser Val Arg Ser Val Gly Lys Ser Ser Gln Arg Ser Phe Thr Thr
20 25 30
10 Ile Thr Gln Leu Arg Ser Thr His Thr Gly Pro Ser Arg Arg Ser Tyr
35 40 45
Leu Ile Trp Trp Asn Gly Gly Pro Lys Arg Thr Ile Ser Tyr Val Ser
50 55 60
15 Arg Arg Phe Arg Ser Phe Arg
65 70

20

(2) INFORMATION FOR SEQ ID NO: 583:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 583:

30 Val Gly Leu Phe Gln Pro Lys Thr Phe Gln Val Pro Val Thr Asp Leu
1 5 10 15
Tyr Ile Phe Ile Lys Ile Tyr Ser Glu Ile Gly Pro Ile Met His Val
20 25 30
35 Leu Cys Pro Gly Tyr Ser Gln Ser Pro Ser Thr Pro Pro Trp Thr
35 40 45

40

(2) INFORMATION FOR SEQ ID NO: 584:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 584:

50 Met Trp Phe Gly Ser Asp Arg Ser Asp Leu Arg Ile Gly Thr Ala Phe
1 5 10 15
Leu Phe Asp Leu Val Cys Asp Leu Cys Ile His Ala Trp Lys Pro Pro
20 25 30
55 Gly Leu Val Arg Phe Ser Phe
35

60

(2) INFORMATION FOR SEQ ID NO: 585:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 585:

Met Leu Asn Thr Ala Ser Leu Asn Leu Pro Trp Lys Val Gln Leu Phe
1 5 10 15

10 Ala His Ala

15 (2) INFORMATION FOR SEQ ID NO: 586:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 586:

Met Ser Ala Cys Leu Leu Leu Phe Leu Ala Phe Ser Trp Lys Arg Lys
1 5 10 15

25 Gly Leu Trp Ser Gly Pro Gly
20

30

(2) INFORMATION FOR SEQ ID NO: 587:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 69 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 587:

Met Leu Pro Pro Phe Ser Leu Val Tyr Thr His Phe Leu Val Ala Ser
40 1 5 10 15

Leu Leu Pro Val Ile Leu Ala Val Phe Pro Asp Ser Ala Gln Ile Val
20 25 30

45 Pro Leu Leu Lys Pro Ile Pro Arg Pro Gln Pro Glu Val Ile Phe Pro
35 40 45

Ser Ser Glu Leu Leu Glu Gln Leu Leu Ser Val Gln Phe Val Trp Gln
50 55 60

50

Ala His Thr Val Ala
65

55

(2) INFORMATION FOR SEQ ID NO: 588:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 77 amino acids

(B) TYPE: amino acid

60

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 588:

5 Met Gly Pro Pro Met Leu Gln Glu Ile Ser Asn Leu Phe Leu Ile Leu
1 5 10 15

Leu Met Met Gly Ala Ile Phe Thr Leu Ala Ala Leu Lys Glu Ser Leu
20 25 30

10 Ser Thr Cys Ile Pro Ala Ile Val Cys Leu Gly Phe Leu Leu Leu Leu
35 40 45

Asn Val Gly Gln Leu Leu Ala Gln Thr Lys Lys Val Val Arg Pro Thr
50 55 60

15 Arg Lys Lys Thr Leu Ser Thr Phe Lys Glu Ser Trp Lys
65 70 75

20

(2) INFORMATION FOR SEQ ID NO: 589:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 155 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 589:

30 Met Ala Leu Leu Leu Ser Val Leu Arg Val Leu Leu Gly Gly Phe Phe
1 5 10 15

Ala Leu Val Gly Leu Ala Lys Leu Ser Glu Glu Ile Ser Ala Pro Val
20 25 30

35 Ser Glu Arg Met Asn Ala Leu Phe Val Gln Phe Ala Glu Val Phe Pro
35 40 45

Leu Lys Val Phe Gly Tyr Gln Pro Asp Pro Leu Asn Tyr Gln Ile Ala
50 55 60

40 Val Gly Phe Leu Glu Leu Leu Ala Gly Leu Leu Val Met Gly Pro
65 70 75 80

45 Pro Met Leu Gln Glu Ile Ser Asn Leu Phe Leu Ile Leu Leu Met Met
85 90 95

Gly Ala Ile Phe Thr Leu Ala Ala Leu Lys Glu Ser Leu Ser Thr Cys
100 105 110

50 Ile Pro Ala Ile Val Cys Leu Gly Phe Leu Leu Leu Leu Asn Val Gly
115 120 125

Gln Leu Leu Ala Gln Thr Lys Lys Val Val Arg Pro Thr Arg Lys Lys
130 135 140

55 Thr Leu Ser Thr Phe Lys Glu Ser Trp Lys Xaa
145 150 155

60

(2) INFORMATION FOR SEQ ID NO: 590:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 590:

5 Met Pro Glu Thr Arg Leu Gly His Arg Gln Gln Phe Ala Val Phe His
10 1 5 10 15
Leu Xaa Pro Val Pro Pro Cys Gly
20

15

(2) INFORMATION FOR SEQ ID NO: 591:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 38 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 591:

25 Met Leu Thr Phe Leu Phe Ser Ala Cys Ala Thr Cys Leu Gly Lys Leu
1 5 10 15
Ala Ser Pro Leu Ala Pro Val Gly Pro Gln Gln Arg Gly Xaa Pro Pro
30 20 25 30
Gly Pro Pro Leu Leu Ser
35

35

(2) INFORMATION FOR SEQ ID NO: 592:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 69 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 592:

45 Met Asp Pro Phe His Tyr Asp Tyr Gln Thr Leu Arg Ile Gly Gly Leu
1 5 10 15
Val Phe Ala Val Val Leu Phe Ser Val Gly Ile Leu Leu Ile Leu Ser
20 25 30
50 Arg Arg Cys Lys Cys Ser Phe Asn Gln Lys Pro Arg Ala Pro Gly Asp
35 40 45
Glu Glu Ala Gln Val Glu Asn Leu Ile Thr Ala Asn Ala Thr Glu Pro
55 50 55 60
Gln Lys Ala Glu Asn
65

60

(2) INFORMATION FOR SEQ ID NO: 593:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 308 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 593:

5 Asn Leu Arg Val Arg Leu Gly Asp Val Ile Ser Ile Gln Pro Cys Pro
 10 1 5 10 15
 Asp Val Lys Tyr Gly Lys Arg Ile His Val Leu Pro Ile Asp Asp Thr
 20 25 30
 15 Val Glu Gly Ile Thr Gly Asn Leu Phe Glu Val Tyr Leu Lys Pro Tyr
 35 40 45
 Phe Leu Glu Ala Tyr Arg Pro Ile Arg Lys Gly Asp Ile Phe Leu Val
 50 55 60
 20 Arg Gly Gly Met Arg Ala Val Glu Phe Lys Val Val Glu Thr Asp Pro
 65 70 75 80
 Ser Pro Tyr Cys Ile Val Ala Pro Asp Thr Val Ile His Cys Glu Gly
 25 85 90 95
 Glu Pro Ile Lys Arg Glu Asp Glu Glu Glu Ser Leu Asn Glu Val Gly
 100 105 110
 30 Tyr Asp Asp Ile Gly Gly Cys Arg Lys Gln Leu Ala Gln Ile Lys Glu
 115 120 125
 Met Val Glu Leu Pro Leu Arg His Pro Ala Leu Phe Lys Ala Ile Gly
 130 135 140
 35 Val Lys Pro Pro Arg Gly Ile Leu Leu Tyr Gly Pro Pro Gly Thr Gly
 145 150 155 160
 Lys Thr Leu Ile Ala Arg Ala Val Ala Asn Glu Thr Gly Ala Phe Phe
 40 165 170 175
 Phe Leu Ile Asn Gly Pro Glu Ile Met Ser Lys Leu Ala Gly Glu Ser
 180 185 190
 45 Glu Ser Asn Leu Arg Lys Ala Phe Glu Glu Ala Glu Lys Asn Ala Pro
 195 200 205
 Ala Ile Ile Phe Ile Asp Glu Leu Asp Ala Ile Ala Pro Lys Arg Glu
 210 215 220
 50 Lys Thr His Gly Glu Val Glu Arg Arg Ile Val Ser Gln Leu Leu Thr
 225 230 235 240
 Leu Met Asp Gly Leu Lys Gln Arg Ala His Val Ile Val Met Ala Ala
 55 245 250 255
 Thr Asn Arg Pro Asn Ser Ile Asp Pro Ala Leu Arg Arg Phe Gly Arg
 260 265 270
 60 Phe Asp Arg Glu Val Asp Ile Gly Ile Pro Asp Ala Thr Gly Arg Leu

275 280 285
Glu Ile Leu Gln Ile His Thr Lys Asn Met Lys Leu Ala Asp Asp Val
290 295 300
5 Asp Leu Glu Gln
305

10

(2) INFORMATION FOR SEQ ID NO: 594:

(i) SEQUENCE CHARACTERISTICS:
15 (A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 594:

20 Met Gln Ile Lys Leu Leu Lys Ser Val Lys Thr Val Phe Ala Ile Thr
1 5 10 15
Leu Leu Val Leu Phe Leu
20

25

(2) INFORMATION FOR SEQ ID NO: 595:

(i) SEQUENCE CHARACTERISTICS:
30 (A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 595:

35 Met Phe Pro Lys Phe Cys Pro Ile Leu Ser Leu Val Asp Phe Ile Ser
1 5 10 15
His Arg Asp Lys Pro Glu Thr Glu
20

40

(2) INFORMATION FOR SEQ ID NO: 596:

45 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 596:

50

Met Leu Ile Glu Cys Ala Trp Gln Leu Met Phe Leu Leu Leu Lys Val
1 5 10 15

55 Glu Gln Leu Gly Ile Leu Asp Lys
20

60

(2) INFORMATION FOR SEQ ID NO: 597:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- 5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 597:
- Met
 1
- 10
- (2) INFORMATION FOR SEQ ID NO: 598:
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 8 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 598:
- 20 Met Cys Ile Met Ser Ala Leu Val
 1 5
- 25 (2) INFORMATION FOR SEQ ID NO: 599:
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 599:
- Met Phe Leu Val Trp Phe Phe Trp Gly Leu Ile Ser Ala Leu Ser Asn
 1 5 10 15
- 35 Val His Thr Pro Ser Arg Leu Pro Ala
 20 25
- 40
- (2) INFORMATION FOR SEQ ID NO: 600:
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 600:
- Met Xaa Gly Leu Ser Leu Ile Leu Thr Val Thr Leu Leu Ala Val Ser
 1 5 10 15
- 50 Asp Ser Ala Ala Thr Cys Ile Val Ala Lys Gly
 20 25
- 55
- (2) INFORMATION FOR SEQ ID NO: 601:
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 61 amino acids
- 60

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 601:

5 Met Trp Thr Arg Ser Ser Arg Cys Leu Leu Leu Cys Ile Pro Gly Xaa
 1 5 10 15

Ser Arg Arg Arg Arg Ala Gly Ser Gly Met Lys Pro Arg Ser Trp Ser
 20 25 30

10 Ala Trp Arg Pro Ser Gly Gly Thr Gly Thr Ser Ser Ser Gln Ser Ser
 35 40 45

15 Thr Gln Ser Arg Thr Leu Ser Ala Thr Ala Ser Pro Ala
 50 55 60

(2) INFORMATION FOR SEQ ID NO: 602:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 602:

Met Arg Glu Thr Ser Ile Arg Val Leu Leu Met Leu Pro Ala Leu Glu
 1 5 10 15

30 Ser Thr Ser Gly Leu Ser Ala Phe Met Gly Leu Gly Thr
 20 25

35

(2) INFORMATION FOR SEQ ID NO: 603:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 69 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 603:

Met Pro Pro Lys Gln Glu Leu Gly Ser Gly Val Gly Glu Leu Ala Lys
 1 5 10 15

45 Asn Ser Lys Arg Gln His Trp Asn His Arg Trp Lys Lys Tyr Leu Lys
 20 25 30

50 Leu Ile Arg Trp Glu Asp Gly Leu Leu Leu Glu Gly Leu Leu Leu Val
 35 40 45

Leu Glu His Cys Ala Thr Met Ala Trp Asp Cys Leu Met Arg Leu Glu
 50 55 60

55 Leu Leu Lys Arg Leu
 65

60

(2) INFORMATION FOR SEQ ID NO: 604:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 604:

Lys Ile Val Tyr Ile Leu Gly Asn Pro Leu Lys Phe Asn Ser Arg Val
1 5 10 15
Ile His His Leu Val Leu Leu Gln
20

(2) INFORMATION FOR SEQ ID NO: 605:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 35 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 605:

Met Asn Leu His Gln Arg Arg Leu Leu Leu Ile Gly His Leu Met Thr
1 5 10 15
Leu Val Lys Ala Ser Lys Ser Phe Ser Phe Thr Glu Ile Thr Ser Ser
20 25 30
Arg Lys Lys
35

(2) INFORMATION FOR SEQ ID NO: 606:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 130 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 606:

Leu Leu Gly Tyr Gly Leu Phe Gly His Cys Ile Val Leu Phe Ile Thr
1 5 10 15
Tyr Asn Ile His Leu His Ala Leu Phe Tyr Leu Phe Trp Leu Leu Val
20 25 30
Gly Gly Leu Ser Thr Leu Arg Met Val Ala Val Leu Val Ser Arg Thr
35 40 45
Val Gly Pro Thr Gln Arg Leu Leu Leu Cys Gly Thr Leu Ala Ala Leu
50 55 60
His Met Leu Phe Leu Leu Tyr Leu His Phe Ala Tyr His Lys Val Xaa
65 70 75 80
Glu Gly Ile Leu Asp Thr Leu Glu Gly Pro Asn Ile Pro Pro Ile Gln
85 90 95

Arg Val Pro Arg Asp Ile Pro Ala Met Leu Pro Ala Ala Arg Leu Pro
100 105 110

5 Thr Thr Val Leu Asn Ala Thr Ala Lys Ala Val Ala Val Thr Leu Gln
115 120 125

Ser His
130

10

(2) INFORMATION FOR SEQ ID NO: 607:

15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 607:

20 Met Leu Val Ile Phe Leu Phe Thr Ser Leu Leu Lys Ile Pro Ser Ser
1 5 10 15

Val Pro Gly Leu Ile Asn Val
20

25

(2) INFORMATION FOR SEQ ID NO: 608:

30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 608:

35

Glu Leu Asp Tyr Ile Leu
1 5

40

(2) INFORMATION FOR SEQ ID NO: 609:

45 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 232 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 609:

50 Met Ala Pro Pro Gly Trp Gln Xaa Xaa Xaa Xaa Xaa Trp Leu Ala Cys
1 5 10 15

Pro Asp Arg Gly Glu Leu Ser Ser Arg Ser Pro Pro Cys Arg Leu Ala
20 25 30

55 Arg Trp Ala Glu Gly Asp Arg Glu Thr Arg Thr Cys Leu Leu Glu Leu
35 40 45

Ser Ala Gln Ser Trp Gly Gly Arg Phe Arg Arg Ser Ser Ala Val Ser
50 55 60

60

Ala Gly Ser Pro Ser Arg Leu His Phe Leu Pro Gln Pro Leu Leu Leu
65 70 75 80

Arg Ser Ser Gly Ile Pro Ala Ala Ala Thr Pro Trp Pro Gln Pro Ala
5 85 90 95

Gly Leu Pro Val Arg Pro Thr Pro Thr Arg Thr Gly Glu Glu Asp Arg
100 105 110

Thr Leu Asp Ile Ser Ile Cys Thr Glu Val Leu Ala Gly Thr Glu Gln
10 115 120 125

Pro Pro Pro Pro Arg Met Thr Ser Pro Ser Ser Ser Pro Val Phe Arg
130 135 140

Leu Glu Thr Leu Asp Gly Gly Gln Glu Asp Gly Ser Glu Ala Asp Arg
15 145 150 155 160

Gly Lys Leu Asp Phe Gly Ser Gly Leu Pro Pro Met Glu Ser Gln Phe
20 165 170 175

Gln Gly Glu Asp Arg Lys Phe Ala Pro Ser Asp Lys Ser Gln Pro Pro
180 185 190

Thr Thr Glu Arg Glu Gln Val Pro Val Ser Arg Ile Gln Thr Asp Leu
25 195 200 205

Thr Glu Ile Gly Ser Ser Met Arg Ser Pro Gly Val Ser Pro Arg Ile
30 210 215 220

Trp Leu Asp Phe Gln Ser Thr Xaa
225 230

35

(2) INFORMATION FOR SEQ ID NO: 610:

(i) SEQUENCE CHARACTERISTICS:

40

- (A) LENGTH: 34 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 610:

Met Val Leu Leu Leu Leu Leu Ala Tyr Val Leu Leu Thr Tyr Ile Leu
45 1 5 10 15

Leu Leu Asn Met Leu Ile Ala Leu Met Xaa Arg Asp Arg Gln Gln Cys
20 25 30

Arg His
50

55

(2) INFORMATION FOR SEQ ID NO: 611:

(i) SEQUENCE CHARACTERISTICS:

60

- (A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 611:

Met Val Phe Glu Gly Phe Ser Ser Ala Phe Cys Leu Ser Ser Thr Ala
1 5 10 15
Pro Thr Ser His Pro
20

10

(2) INFORMATION FOR SEQ ID NO: 612:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 612:

Gly Lys Lys Asn Gln Leu Leu Val Ile
1 5
20

(2) INFORMATION FOR SEQ ID NO: 613:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 613:

Met Val Trp Val Leu Trp Ser Ala Pro Ser Leu Ala Pro Pro Trp Val
1 5 10 15

Gly Pro Cys Trp Pro Ser Thr Gly Asn Cys Cys Leu Cys
20 25
35

(2) INFORMATION FOR SEQ ID NO: 614:

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 614:

Met Ala Lys Arg Ser Pro Gly Gly Cys Gly Ser Gly Leu Ile Leu Leu
1 5 10 15

50

Cys Cys Gln Pro Cys Arg Pro Thr Ser Ser Ala Pro Met Arg
20 25 30

55

(2) INFORMATION FOR SEQ ID NO: 615:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 113 amino acids

(B) TYPE: amino acid

60

(2) INFORMATION FOR SEQ ID NO: 618:

(i) SEQUENCE CHARACTERISTICS:

5

(A) LENGTH: 52 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 618:

10 Gln Ala Asp Asp Leu Gln Ala Thr Val Ala Ala Leu Cys Val Leu Arg
 1 5 10 15

Gly Gly Gly Pro Trp Ala Gly Ser Trp Leu Ser Pro Lys Thr Pro Gly
 20 25 30

15 Ala Met Gly Gly Asp Leu Val Leu Gly Leu Gly Ala Leu Arg Arg Arg
 35 40 45

Lys Arg Leu Leu
 20 50

(2) INFORMATION FOR SEQ ID NO: 619:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 232 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 619:

Glu Gln Glu Lys Ser Leu Ala Gly Trp Ala Leu Val Leu Ala Xaa Xaa
 1 5 10 15

35 Gly Ile Gly Leu Met Val Leu His Ala Glu Met Leu Trp Phe Gly Gly
 20 25 30

Cys Ser Ala Val Asn Ala Thr Gly His Leu Ser Asp Thr Leu Trp Leu
 35 40 45

40 Ile Pro Ile Thr Phe Leu Thr Ile Gly Tyr Gly Asp Val Val Pro Gly
 50 55 60

45 Thr Met Trp Gly Lys Ile Val Cys Leu Cys Thr Gly Val Met Gly Val
 65 70 75 80

Cys Cys Thr Ala Leu Leu Val Ala Val Val Ala Arg Lys Leu Glu Phe
 85 90 95

50 Asn Lys Ala Glu Lys His Val His Asn Phe Met Met Asp Ile Gln Tyr
 100 105 110

Thr Lys Glu Met Lys Glu Ser Ala Ala Arg Val Leu Gln Glu Ala Trp
 115 120 125

55 Met Phe Tyr Lys His Thr Arg Arg Lys Glu Ser His Ala Ala Arg Xaa
 130 135 140

60 His Gln Arg Xaa Leu Leu Ala Ala Ile Asn Ala Phe Arg Gln Val Arg
 145 150 155 160

Leu Lys His Arg Lys Leu Arg Glu Gln Val Asn Ser Met Val Asp Ile
 165 170 175
 5 Ser Lys Met His Met Ile Leu Tyr Asp Leu Gln Gln Asn Leu Ser Ser
 180 185 190
 Ser His Arg Ala Leu Glu Lys Gln Ile Asp Thr Leu Ala Gly Lys Leu
 195 200 205
 10 Asp Ala Leu Thr Glu Leu Leu Ser Thr Ala Leu Gly Pro Arg Gln Leu
 210 215 220
 Pro Glu Pro Ser Gln Gln Ser Lys
 15 225 230

20 (2) INFORMATION FOR SEQ ID NO: 620:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 36 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 620:

Tyr Gln Ala His His Val Ser Arg Asn Lys Arg Gly Gln Val Val Gly
 1 5 10 15
 30 Thr Arg Gly Gly Phe Arg Gly Cys Thr Val Trp Leu Thr Gly Leu Ser
 20 25 30
 Gly Ala Gly Lys
 35

40 (2) INFORMATION FOR SEQ ID NO: 621:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 57 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 621:

Leu Gln Cys Glu Ile Cys Gly Phe Thr Cys Arg Gln Lys Ala Ser Leu
 1 5 10 15
 50 Asn Trp His Met Lys Lys His Asp Ala Asp Ser Phe Tyr Gln Phe Ser
 20 25 30
 Cys Asn Ile Cys Gly Lys Lys Phe Glu Lys Lys Asp Ser Val Val Ala
 35 40 45
 55 His Lys Ala Lys Ser His Pro Glu Val
 50 55

60 (2) INFORMATION FOR SEQ ID NO: 622:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 622:

Ile Thr Ser Thr Asp Ile Leu Gly Thr Asn Pro Glu Ser Leu Thr Gln
1 5 10 15

Pro Ser Asp

(2) INFORMATION FOR SEQ ID NO: 623:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 623:

Asn Ser Thr Ser Gly Glu Cys Leu Leu Leu Glu Ala Glu Gly Met Ser
1 5 10 15

Lys Ser Tyr

(2) INFORMATION FOR SEQ ID NO: 624:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 51 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 624:

Cys Ser Gly Thr Glu Arg Val Ser Leu Met Ala Asp Gly Lys Ile Phe
1 5 10 15

Val Gly Ser Gly Ser Ser Gly Gly Thr Glu Gly Leu Val Met Asn Ser
20 25 30

Asp Ile Leu Gly Ala Thr Thr Glu Val Leu Ile Glu Asp Ser Asp Ser
35 40 45

Ala Gly Pro
50

(2) INFORMATION FOR SEQ ID NO: 625:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 625:

Ile Gln Tyr Val Arg Cys Glu Met Glu Gly Cys Gly Thr Val Leu Ala
 1 5 10 15
 5 His Pro Arg Tyr Leu Gln His His Ile Lys Tyr Gln His Leu Leu Lys
 20 25 30
 Lys Lys Tyr Val Cys Pro His Pro Ser Cys Gly Arg Leu Phe Arg Leu
 35 40 45
 10 Gln Lys Gln Leu Leu Arg His Ala Lys His His Thr
 50 55 60

15

(2) INFORMATION FOR SEQ ID NO: 626:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 626:

20 Asp Gln Arg Asp Tyr Ile Cys Glu Tyr Cys Ala Arg Ala Phe Lys Ser
 25 1 5 10 15
 Ser His Asn Leu Ala Val His Arg Met Ile His Thr Gly Glu Lys
 20 25 30

30

(2) INFORMATION FOR SEQ ID NO: 627:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 627:

35 Arg Ser Ser Arg Ser Lys Thr Gly Ser Leu Gln Leu Ile Cys Lys Ser
 40 1 5 10 15
 Glu Pro Asn Thr Asp Gln Leu Asp Tyr
 20 25

45

(2) INFORMATION FOR SEQ ID NO: 628:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 183 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 628:

55

Leu Gln Cys Glu Ile Cys Gly Phe Thr Cys Arg Gln Lys Ala Ser Leu
 1 5 10 15
 Asn Trp His Met Lys Lys His Asp Ala Asp Ser Phe Tyr Gln Phe Ser
 20 25 30

60

Cys Asn Ile Cys Gly Lys Lys Phe Glu Lys Lys Asp Ser Val Val Ala
35 40 45

5 His Lys Ala Lys Ser His Pro Glu Val Xaa Ile Thr Ser Thr Asp Ile
50 55 60

Leu Gly Thr Asn Pro Glu Ser Leu Thr Gln Pro Ser Asp Xaa Asn Ser
65 70 75 80

10 Thr Ser Gly Glu Cys Leu Leu Leu Glu Ala Glu Gly Met Ser Lys Ser
85 90 95

Tyr Xaa Cys Ser Gly Thr Glu Arg Val Ser Leu Met Ala Asp Gly Lys
15 100 105 110

Ile Phe Val Gly Ser Gly Ser Ser Gly Gly Thr Glu Gly Leu Val Met
115 120 125

20 Asn Ser Asp Ile Leu Gly Ala Thr Thr Glu Val Leu Ile Glu Asp Ser
130 135 140

Asp Ser Ala Gly Pro Xaa Gln Arg Asp Tyr Ile Cys Glu Tyr Cys Ala
145 150 155 160

25 Arg Ala Phe Lys Ser Ser His Asn Leu Ala Val His Arg Met Ile His
165 170 175

Thr Gly Glu Lys His Tyr Xaa
30 180

35 (2) INFORMATION FOR SEQ ID NO: 629:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 629:

Gln Tyr Val Arg Cys Glu Met Glu Gly Cys Gly Thr Val Leu Ala His
1 5 10 15

45 Pro Arg Tyr Leu Gln His His Ile Lys Tyr Gln His Leu Leu Lys Lys
20 25 30

Lys Tyr Val Cys Pro His Pro Ser Cys Gly Arg Leu Phe Arg Leu Gln
35 40 45

50 Lys Gln Leu Leu Arg His Ala Lys His His Thr Asp
50 55 60

55 (2) INFORMATION FOR SEQ ID NO: 630:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid

60

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEO ID NO: 630:

Pro Phe Lys Asp Asp Pro Arg Asp Glu Thr Tyr Lys Pro His Leu Glu
5 1 5 10 15

Arg Glu Thr Pro Lys Pro Arg Arg Lys Ser Gly
20 25

10

(2) INFORMATION FOR SEQ ID NO: 631:

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 110 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 631:

20 Glu Met Phe Asp Ser Leu Ser Tyr Phe Lys Gly Ser Ser Leu Leu Leu
 1 5 10 15

Met Leu Lys Thr Tyr Leu Ser Glu Asp Val Phe Gln His Ala Val Val
20 25 30

25

Leu Tyr Leu His Asn His Ser Tyr Ala Ser Ile Gln Ser Asp Asp Leu
35 40 45

30 Trp Asp Ser Phe Asn Glu Val Thr Asn Gln Thr Leu Asp Val Lys Arg
50 55 60

Met Met Lys Thr Trp Thr Leu Gln Lys Gly Phe Pro Leu Val Thr Val
65 70 75 80

35 Gln Lys Lys Gly Lys Glu Leu Phe Ile Gln Gln Glu Arg Phe Phe Leu
85 90 95

Asn Met Lys Pro Glu Ile Gln Pro Ser Asp Thr Arg Tyr Met
100 105 110

40

(2) INFORMATION FOR SEQ ID NO: 632:

45 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 632:

50

Leu Glu Lys Val Ala Ser Val Gly Asn Ser Arg Pro Thr Gly Gln Gln
1 5 10 15

Leu Glu Ser Leu Gly Leu Leu Ala
20

55

(2) INFORMATION FOR SEQ ID NO: 633:

60

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 633:

Val His Arg Glu Glu Ala Ser Cys Tyr Cys Gln Ala Glu Pro Ser Gly
1 5 10 15

10 Asp Leu

15 (2) INFORMATION FOR SEQ ID NO: 634:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 634:

Arg Pro Ala Leu Arg Gln Ala Gly Gly Gly Thr Arg Glu Pro Arg Gln
1 5 10 15

25 Lys Arg Trp Ala Gly Leu
20

30

(2) INFORMATION FOR SEQ ID NO: 635:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 635:

40 Ala Val Asn Phe Arg Pro Gln Arg Ser Gln Ser Met
1 5 10

45 (2) INFORMATION FOR SEQ ID NO: 636:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 636:

Met Ile Thr Asp Val Gln Leu Ala Ile Phe Ala Asn Met Leu Gly Val
1 5 10 15

55 Ser Leu Phe Leu Leu Val Val Leu Tyr His Tyr Val Ala Val Asn Asn
20 25 30Pro Lys Lys Gln Glu
35

60

(2) INFORMATION FOR SEQ ID NO: 637:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 342 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 637:

Glu Glu Met Ala Asp Ser Val Lys Thr Phe Leu Gln Asp Leu Ala Arg
 1 5 10 15

15

Gly Ile Lys Asp Ser Ile Trp Gly Ile Cys Thr Ile Ser Lys Leu Asp
 20 25 30

Ala Arg Ile Gln Gln Lys Arg Glu Glu Gln Arg Arg Arg Ala Ser
 35 40 45

20

Ser Val Leu Ala Gln Arg Arg Ala Gln Ser Ile Glu Arg Lys Gln Glu
 50 55 60

Ser Glu Pro Arg Ile Val Ser Arg Ile Phe Gln Cys Cys Ala Trp Asn
 65 70 75 80

25

Gly Gly Val Phe Trp Phe Ser Leu Leu Leu Phe Tyr Arg Val Phe Ile
 85 90 95

30

Pro Val Leu Gln Ser Val Thr Ala Arg Ile Ile Gly Asp Pro Ser Leu
 100 105 110

His Gly Asp Val Trp Ser Trp Leu Glu Phe Phe Leu Thr Ser Ile Phe
 115 120 125

35

Ser Ala Leu Trp Val Leu Pro Leu Phe Val Leu Ser Lys Val Val Asn
 130 135 140

Ala Ile Trp Phe Gln Asp Ile Ala Asp Leu Ala Phe Glu Val Ser Gly
 145 150 155 160

40

Arg Lys Pro His Pro Phe Pro Ser Val Ser Lys Ile Ile Ala Asp Met
 165 170 175

Leu Phe Asn Leu Leu Leu Gln Ala Leu Phe Leu Ile Gln Gly Met Phe
 180 185 190

45

Val Ser Leu Phe Pro Ile His Leu Val Gly Gln Leu Val Ser Leu Leu
 195 200 205

50

His Met Ser Leu Leu Tyr Ser Leu Tyr Cys Phe Glu Tyr Arg Trp Phe
 210 215 220

Asn Lys Gly Ile Glu Met His Gln Arg Leu Ser Asn Ile Glu Arg Asn
 225 230 235 240

55

Trp Pro Tyr Tyr Phe Gly Phe Gly Leu Pro Leu Ala Phe Leu Thr Ala
 245 250 255

60

Met Gln Ser Ser Tyr Ile Ile Ser Gly Cys Leu Phe Ser Ile Leu Phe
 260 265 270

Pro Leu Phe Ile Ile Ser Ala Asn Glu Ala Lys Thr Pro Gly Lys Ala
 275 280 285
 5 Tyr Leu Phe Gln Leu Arg Leu Phe Ser Leu Val Val Phe Leu Ser Asn
 290 295 300
 Arg Leu Phe His Lys Thr Val Tyr Leu Gln Ser Ala Leu Ser Ser Ser
 305 310 315 320
 10 Thr Ser Ala Glu Lys Phe Pro Ser Pro His Pro Ser Pro Ala Lys Leu
 325 330 335
 Lys Ala Thr Ala Gly His
 15 340

(2) INFORMATION FOR SEQ ID NO: 638:
 20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 529 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 638:

Met Ala Lys Phe Met Thr Pro Val Ile Gln Asp Asn Pro Ser Gly Trp
 1 5 10 15
 30 Gly Pro Cys Ala Val Pro Glu Gln Phe Arg Asp Met Pro Tyr Gln Pro
 20 25 30
 Phe Ser Lys Gly Asp Arg Leu Gly Lys Val Ala Asp Trp Thr Gly Ala
 35 35 40 45
 Thr Tyr Gln Asp Lys Arg Tyr Thr Asn Lys Tyr Ser Ser Gln Phe Gly
 50 55 60
 40 Gly Gly Ser Gln Tyr Ala Tyr Phe His Glu Glu Asp Glu Ser Ser Phe
 65 70 75 80
 Gln Leu Val Asp Thr Ala Arg Thr Gln Lys Thr Ala Tyr Gln Arg Asn
 85 90 95
 45 Arg Met Arg Phe Ala Gln Arg Asn Leu Arg Arg Asp Lys Asp Arg Arg
 100 105 110
 Asn Met Leu Gln Phe Asn Leu Gln Ile Leu Pro Lys Ser Ala Lys Gln
 115 120 125
 50 Lys Glu Arg Glu Arg Ile Arg Leu Gln Lys Lys Phe Gln Lys Gln Phe
 130 135 140
 Gly Val Arg Gln Lys Trp Asp Gln Lys Ser Gln Lys Pro Arg Asp Ser
 145 150 155 160
 Ser Val Glu Val Arg Ser Asp Trp Glu Val Lys Glu Glu Met Asp Phe
 165 170 175
 60 Pro Gln Leu Met Lys Met Arg Tyr Leu Glu Val Ser Glu Pro Gln Asp

	180	185	190
	Ile Glu Cys Cys Gly Ala Leu Glu Tyr Tyr Asp Lys Ala Phe Asp Arg		
	195	200	205
5	Ile Thr Thr Arg Ser Glu Lys Pro Leu Arg Xaa Xaa Lys Arg Ile Phe		
	210	215	220
10	His Thr Val Thr Thr Thr Asp Asp Pro Val Ile Arg Lys Leu Ala Lys		
	225	230	235
	Thr Gln Gly Asn Val Phe Ala Thr Asp Ala Ile Leu Ala Thr Leu Met		
	245	250	255
15	Ser Cys Thr Arg Ser Val Tyr Ser Trp Asp Ile Val Val Gln Arg Val		
	260	265	270
	Gly Ser Lys Leu Phe Phe Asp Lys Arg Asp Asn Ser Asp Phe Asp Leu		
	275	280	285
20	Leu Thr Val Ser Glu Thr Ala Asn Glu Pro Pro Gln Asp Glu Gly Asn		
	290	295	300
25	Ser Phe Asn Ser Pro Arg Asn Leu Ala Met Glu Ala Thr Tyr Ile Asn		
	305	310	315
	His Asn Phe Ser Gln Gln Cys Leu Arg Met Gly Lys Glu Arg Tyr Asn		
	325	330	335
30	Phe Pro Asn Pro Asn Pro Phe Val Glu Asp Asp Met Asp Lys Asn Glu		
	340	345	350
	Ile Ala Ser Val Ala Tyr Arg Tyr Arg Ser Gly Lys Leu Gly Asp Asp		
	355	360	365
35	Ile Asp Leu Ile Val Arg Cys Glu His Asp Gly Val Met Thr Gly Ala		
	370	375	380
40	Asn Gly Glu Val Ser Phe Ile Asn Ile Lys Thr Leu Asn Glu Trp Asp		
	385	390	395
	Ser Arg His Cys Asn Gly Val Asp Trp Arg Gln Lys Leu Asp Ser Gln		
	405	410	415
45	Arg Gly Ala Val Ile Ala Thr Glu Leu Lys Asn Asn Ser Tyr Lys Leu		
	420	425	430
	Ala Arg Trp Thr Cys Cys Ala Leu Leu Ala Gly Ser Glu Tyr Leu Lys		
	435	440	445
50	Leu Gly Tyr Val Ser Arg Tyr His Val Lys Asp Ser Ser Arg His Val		
	450	455	460
55	Ile Leu Gly Thr Gln Gln Phe Lys Pro Asn Glu Phe Ala Ser Gln Ile		
	465	470	475
	Asn Leu Ser Val Glu Asn Ala Trp Gly Ile Leu Arg Cys Val Ile Asp		
	485	490	495
60	Ile Cys Met Lys Leu Glu Glu Gly Lys Tyr Leu Ile Leu Lys Asp Pro		

500 505 510

Asn Lys Gln Val Ile Arg Val Tyr Ser Leu Pro Asp Gly Thr Phe Ser
 515 520 525

5 Ser

10 (2) INFORMATION FOR SEQ ID NO: 639:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 194 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 639:

20 Lys Lys Arg His Thr Asp Val Gln Phe Tyr Thr Glu Val Gly Glu Ile
 1 5 10 15

Thr Thr Asp Leu Gly Lys His Gln His Met His Asp Arg Asp Asp Leu
 20 25 30

25 Tyr Ala Glu Gln Met Glu Arg Glu Met Arg His Lys Leu Lys Thr Ala
 35 40 45

Phe Lys Asn Phe Ile Glu Lys Val Glu Ala Leu Thr Lys Glu Glu Leu
 50 55 60

30 Glu Phe Glu Val Pro Phe Arg Asp Leu Gly Phe Asn Gly Ala Pro Tyr
 65 70 75 80

35 Arg Ser Thr Cys Leu Leu Gln Pro Thr Ser Ser Ala Leu Val Asn Ala
 85 90 95

Thr Glu Trp Pro Pro Phe Val Val Thr Leu Asp Glu Val Glu Leu Ile
 100 105 110

40 His Phe Xaa Arg Val Gln Phe His Leu Lys Asn Phe Asp Met Val Ile
 115 120 125

Val Tyr Lys Asp Tyr Ser Lys Lys Val Thr Met Ile Asn Ala Ile Pro
 130 135 140

45 Val Ala Ser Leu Asp Pro Ile Lys Glu Trp Leu Asn Ser Cys Asp Leu
 145 150 155 160

50 Lys Tyr Thr Glu Gly Val Gln Ser Leu Asn Trp Thr Lys Ile Met Lys
 165 170 175

Thr Ile Val Asp Asp Pro Glu Gly Phe Phe Glu Gln Gly Gly Trp Ser
 180 185 190

55 Phe Leu

60 (2) INFORMATION FOR SEQ ID NO: 640:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 70 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 640:

Arg Ser Gly Leu Gly Leu Gly Ile Thr Ile Ala Phe Leu Ala Thr Leu
1 5 10 15
Ile Thr Gln Phe Leu Val Tyr Asn Gly Val Tyr Gln Tyr Thr Ser Pro
20 25 30
Asp Phe Leu Tyr Ile Arg Ser Trp Leu Pro Cys Ile Phe Phe Ser Gly
35 40 45
Gly Val Thr Val Gly Asn Ile Gly Arg Gln Leu Ala Met Gly Val Pro
50 55 60
Glu Lys Pro His Ser Asp
65 70

(2) INFORMATION FOR SEQ ID NO: 641:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 101 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 641:

Val Thr Gln Pro Lys His Leu Ser Ala Ser Met Gly Gly Ser Val Glu
1 5 10 15
Ile Pro Phe Ser Phe Tyr Tyr Pro Trp Glu Leu Ala Xaa Xaa Pro Xaa
20 25 30
Val Arg Ile Ser Trp Arg Arg Gly His Phe His Gly Gln Ser Phe Tyr
35 40 45
Ser Thr Arg Pro Pro Ser Ile His Lys Asp Tyr Val Asn Arg Leu Phe
50 55 60
Leu Asn Trp Thr Glu Gly Gln Glu Ser Gly Phe Leu Arg Ile Ser Asn
65 70 75 80
Leu Arg Lys Glu Asp Gln Ser Val Tyr Phe Cys Arg Val Glu Leu Asp
85 90 95
Thr Arg Arg Ser Gly
100

(2) INFORMATION FOR SEQ ID NO: 642:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 233 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 642:

5 Met Glu Ala Gln Gln Val Asn Glu Ala Glu Ser Ala Arg Glu Gln Leu
 1 5 10 15
 Gln Xaa Leu His Asp Gln Ile Ala Gly Gln Lys Ala Ser Lys Gln Glu
 20 25 30
 10 Leu Glu Thr Glu Leu Glu Arg Leu Lys Gln Glu Phe His Tyr Ile Glu
 35 40 45
 Glu Asp Leu Tyr Arg Thr Lys Asn Thr Leu Gln Ser Arg Ile Lys Asp
 50 55 60
 15 Arg Asp Glu Glu Ile Gln Lys Leu Arg Asn Gln Leu Thr Asn Lys Thr
 65 70 75 80
 Leu Ser Asn Ser Ser Gln Ser Glu Leu Glu Asn Arg Leu His Gln Leu
 20 85 90 95
 Thr Glu Thr Leu Ile Gln Lys Gln Thr Met Leu Glu Ser Leu Ser Thr
 100 105 110
 25 Glu Lys Asn Ser Leu Val Phe Gln Leu Glu Arg Leu Glu Gln Gln Met
 115 120 125
 Asn Ser Ala Ser Gly Ser Ser Ser Asn Gly Ser Ser Ile Asn Met Ser
 130 135 140
 Gly Ile Asp Asn Gly Glu Gly Thr Arg Leu Arg Asn Val Pro Val Leu
 145 150 155 160
 35 Phe Asn Asp Thr Glu Thr Asn Leu Ala Gly Met Tyr Gly Lys Val Arg
 165 170 175
 Lys Ala Ala Ser Ser Ile Asp Gln Phe Ser Ile Arg Leu Gly Ile Phe
 180 185 190
 40 Leu Arg Arg Tyr Pro Ile Ala Arg Val Phe Val Ile Ile Tyr Met Ala
 195 200 205
 Leu Leu His Leu Trp Val Met Ile Val Leu Leu Thr Tyr Thr Pro Glu
 210 215 220
 45 Met His His Asp Gln Pro Tyr Gly Lys
 225 230

50

(2) INFORMATION FOR SEQ ID NO: 643:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 643:

60 Ile Arg His Glu Gln His Pro Asn Phe Ser Leu Glu Met His Ser Lys
 1 5 10 15

Gly Ser Ser Leu Leu Leu Phe Leu Pro Gln Leu Ile Leu Ile Leu Pro
20 25 30

5 Val Cys Ala His Leu His Glu Glu Leu Asn Cys
35 40

10 (2) INFORMATION FOR SEQ ID NO: 644:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 63 amino acids

(B) TYPE: amino acid

15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 644:

Ser Phe Phe Ile Ser Glu Glu Lys Gly His Leu Leu Leu Gln Ala Glu
1 5 10 15

20 Arg His Pro Trp Val Ala Gly Ala Leu Val Gly Val Ser Gly Gly Leu
20 25 30

25 Thr Leu Thr Thr Cys Ser Gly Pro Thr Glu Lys Pro Ala Thr Lys Asn
35 40 45

Tyr Phe Leu Lys Arg Leu Leu Gln Glu Met His Ile Arg Ala Asn
50 55 60

30

35

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>116</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97897
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<div style="text-align: center; border-bottom: 1px solid black; margin-bottom: 5px;">For receiving Office use only</div> <div style="padding: 5px;"><input checked="checked" type="checkbox"/> This sheet was received with the international application</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer <div style="text-align: center;">Susan White PCT International Division</div></div>	<div style="text-align: center; border-bottom: 1px solid black; margin-bottom: 5px;">For International Bureau use only</div> <div style="padding: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer</div>
---	---

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>116</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 209043
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<div style="border-bottom: 1px solid black; margin-bottom: 5px;">For receiving Office use only</div> <div style="padding: 5px;"><input checked="checked" type="checkbox"/> This sheet was received with the international application</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer <div style="text-align: center;">Susan White PCT International Division</div></div>	<div style="border-bottom: 1px solid black; margin-bottom: 5px;">For International Bureau use only</div> <div style="padding: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer</div>
--	---

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>119</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit September 4, 1997	Accession Number 209235
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
<div style="text-align: center;">For receiving Office use only</div> <div style="border: 1px solid black; padding: 5px; margin-top: 5px;"><input checked="checked" type="checkbox"/> This sheet was received with the international application</div> <div style="border: 1px solid black; padding: 5px; margin-top: 5px;">Authorized officer <div style="text-align: center;">Susan White PCT International Division</div></div>	<div style="text-align: center;">For International Bureau use only</div> <div style="border: 1px solid black; padding: 5px; margin-top: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border: 1px solid black; padding: 5px; margin-top: 5px;">Authorized officer</div>

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>122</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97898
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<div style="text-align: center; border-bottom: 1px solid black; margin-bottom: 5px;">For receiving Office use only</div> <div style="padding: 5px;"><input checked="" type="checkbox"/> This sheet was received with the international application</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer <div style="text-align: center;">Susan White PCT International Division</div></div>	<div style="text-align: center; border-bottom: 1px solid black; margin-bottom: 5px;">For International Bureau use only</div> <div style="padding: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer</div>
--	---

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>122</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 209044
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
<div style="text-align: center;">For receiving Office use only</div> <div style="border: 1px solid black; padding: 5px; margin-top: 5px;"><input checked="checked" type="checkbox"/> This sheet was received with the international application</div> <div style="border: 1px solid black; padding: 5px; margin-top: 5px;">Authorized officer <div style="text-align: center;">Susan White PCT International Division</div></div>	<div style="text-align: center;">For International Bureau use only</div> <div style="border: 1px solid black; padding: 5px; margin-top: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border: 1px solid black; padding: 5px; margin-top: 5px;">Authorized officer</div>

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>126</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97899
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<div style="border-bottom: 1px solid black; margin-bottom: 5px;">For receiving Office use only</div> <div style="display: flex; align-items: center; margin-bottom: 10px;"><input checked="checked" type="checkbox"/> This sheet was received with the international application</div> <div style="display: flex; align-items: center;"><div style="flex: 1;">Authorized officer</div><div style="text-align: center; margin-left: 20px;">Susan White PCT International Division</div></div>	<div style="border-bottom: 1px solid black; margin-bottom: 5px;">For International Bureau use only</div> <div style="display: flex; align-items: center; margin-bottom: 10px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="display: flex; align-items: center;"><div style="flex: 1;">Authorized officer</div><div style="flex: 1;"></div></div>
--	---

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>126</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 209045
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<div style="text-align: center;">For receiving Office use only</div> <div style="padding: 5px;"><input checked="checked" type="checkbox"/> This sheet was received with the international application</div> <div style="padding: 5px;">Authorized officer <div style="text-align: center;">Susan White POT International Division</div></div>	<div style="text-align: center;">For International Bureau use only</div> <div style="padding: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="padding: 5px;">Authorized officer</div>
--	--

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>130</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit April 28, 1997	Accession Number 209011
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<div style="text-align: center; border-top: 1px solid black; border-bottom: 1px solid black;">For receiving Office use only</div> <div style="padding: 5px;"><input checked="" type="checkbox"/> This sheet was received with the international application</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer <div style="text-align: center;">Susan White PCT International Division</div></div>	<div style="text-align: center; border-top: 1px solid black; border-bottom: 1px solid black;">For International Bureau use only</div> <div style="padding: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer</div>
---	--

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>131</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97900
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit") 	

For receiving Office use only	For International Bureau use only
<input checked="checked" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer Susan White PCT International Division	Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>137</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97901
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<div style="border-bottom: 1px solid black; margin-bottom: 5px;">For receiving Office use only</div> <div style="display: flex; align-items: center; margin-bottom: 10px;"><input checked="checked" type="checkbox"/> This sheet was received with the international application</div> <div style="border-bottom: 1px solid black; margin-bottom: 10px;">Authorized officer</div> <div style="text-align: center;">Susan White PCT International Division</div>	<div style="border-bottom: 1px solid black; margin-bottom: 5px;">For International Bureau use only</div> <div style="display: flex; align-items: center; margin-bottom: 10px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border-bottom: 1px solid black; margin-bottom: 10px;">Authorized officer</div>
---	--

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>131</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 209046
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<div style="border-bottom: 1px solid black; margin-bottom: 5px;">For receiving Office use only</div> <div style="padding: 5px;"><input checked="checked" type="checkbox"/> This sheet was received with the international application</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer <div style="text-align: center;">Susan White PCT International Division</div></div>	<div style="border-bottom: 1px solid black; margin-bottom: 5px;">For International Bureau use only</div> <div style="padding: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer</div>
---	---

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>137</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 209047
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
<div style="text-align: center;">For receiving Office use only</div> <div style="border: 1px solid black; padding: 5px; margin-top: 5px;"><input checked="checked" type="checkbox"/> This sheet was received with the international application</div> <div style="border: 1px solid black; padding: 5px; margin-top: 5px;">Authorized officer Susan White PCT International Division</div>	<div style="text-align: center;">For International Bureau use only</div> <div style="border: 1px solid black; padding: 5px; margin-top: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border: 1px solid black; padding: 5px; margin-top: 5px;">Authorized officer</div>

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>137</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 22, 1997	Accession Number 209076
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<div style="text-align: center; border-bottom: 1px solid black; margin-bottom: 5px;">For receiving Office use only</div> <div style="padding: 5px;"><input checked="checked" type="checkbox"/> This sheet was received with the international application</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer <div style="text-align: center;">Susan White PCT International Division</div></div>	<div style="text-align: center; border-bottom: 1px solid black; margin-bottom: 5px;">For International Bureau use only</div> <div style="padding: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer</div>
---	---

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>140</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 .. United States of America	
Date of deposit August 21, 1997	Accession Number 209215
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<div style="border-bottom: 1px solid black; margin-bottom: 5px;">For receiving Office use only</div> <div style="padding: 5px;"><input checked="" type="checkbox"/> This sheet was received with the international application</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer <div style="text-align: center;">Susan White PCT International Division</div></div>	<div style="border-bottom: 1px solid black; margin-bottom: 5px;">For International Bureau use only</div> <div style="padding: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer</div>
---	---

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>160</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <p style="text-align: center;">American Type Culture Collection</p>	
Address of depositary institution (including postal code and country) <p>12301 Parklawn Drive Rockville, Maryland 20852 United States of America</p>	
Date of deposit February 26, 1997	Accession Number 97904
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	For International Bureau use only
<input checked="checked" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer Susan White PCT International Division	Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>154</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>July 3, 1997</u>	Accession Number <u>209139</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit") 	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer Susan White PCT International Division	Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>153</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 209049
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer Susan White PCT International Division	Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>153</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <p style="text-align: center;">American Type Culture Collection</p>	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97903
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	For International Bureau use only
<input checked="checked" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer <p style="text-align: center;">Susan White PCT International Division</p>	Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>142</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit <u>June 12, 1997</u>	Accession Number <u>209119</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<div style="text-align: center; border-bottom: 1px solid black;">For receiving Office use only</div> <div style="padding: 5px;"><input checked="" type="checkbox"/> This sheet was received with the international application</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer <div style="text-align: center;">Susan White PCT International Division</div></div>	<div style="text-align: center; border-bottom: 1px solid black;">For International Bureau use only</div> <div style="padding: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer</div>
--	---

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>146</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97902
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<div style="text-align: center; border-bottom: 1px solid black; margin-bottom: 5px;">For receiving Office use only</div> <div style="padding: 5px;"><input checked="checked" type="checkbox"/> This sheet was received with the international application</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer Susan White PCT International Division</div>	<div style="text-align: center; border-bottom: 1px solid black; margin-bottom: 5px;">For International Bureau use only</div> <div style="padding: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border-top: 1px solid black; padding: 5px;">Authorized officer</div>
---	---

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>146</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <p style="text-align: center;">American Type Culture Collection</p>	
Address of depositary institution (including postal code and country) <p>12301 Parklawn Drive Rockville, Maryland 20852 United States of America</p>	
Date of deposit <u>May 15, 1997</u>	Accession Number <u>209048</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit") 	

For receiving Office use only



This sheet was received with the international application

Authorized officer

Susan White
PCT International Division

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>160</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 209050
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer Susan White PCT International Division	Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>142</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>February 12, 1998</u>	Accession Number <u>209627</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<div style="border-bottom: 1px solid black; margin-bottom: 5px;">For receiving Office use only</div> <div style="display: flex; align-items: center; margin-bottom: 10px;"><input checked="" type="checkbox"/> This sheet was received with the international application</div> <div style="display: flex; justify-content: space-between; margin-top: 10px;"><div style="width: 40%;">Authorized officer</div><div style="width: 60%;">Susan White PCT International Division</div></div>	<div style="border-bottom: 1px solid black; margin-bottom: 5px;">For International Bureau use only</div> <div style="display: flex; align-items: center; margin-bottom: 10px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="display: flex; justify-content: space-between; margin-top: 10px;"><div style="width: 40%;">Authorized officer</div><div style="width: 60%;"></div></div>
--	--

What Is Claimed Is:

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:
- 5 (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- 10 (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
- 15 (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X, having biological activity;
- 20 (f) a polynucleotide which is a variant of SEQ ID NO:X;
- (g) a polynucleotide which is an allelic variant of SEQ ID NO:X;
- (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;
- (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.
- 25
2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.
- 30
3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.
- 35

4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

5

5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

10

6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

15

7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.

8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.

20

9. A recombinant host cell produced by the method of claim 8.

10. The recombinant host cell of claim 9 comprising vector sequences.

25

11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;

30

(c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

35

(e) a secreted form of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

- (g) a variant of SEQ ID NO:Y;
- (h) an allelic variant of SEQ ID NO:Y; or
- (i) a species homologue of the SEQ ID NO:Y.

12. The isolated polypeptide of claim 11, wherein the secreted form or the
5 full length protein comprises sequential amino acid deletions from either the C-terminus
or the N-terminus.

13. An isolated antibody that binds specifically to the isolated polypeptide of
claim 11.

10

14. A recombinant host cell that expresses the isolated polypeptide of claim
11.

15. A method of making an isolated polypeptide comprising:

- 15
- (a) culturing the recombinant host cell of claim 14 under conditions such that
said polypeptide is expressed; and
 - (b) recovering said polypeptide.

16. The polypeptide produced by claim 15.

20

17. A method for preventing, treating, or ameliorating a medical condition,
comprising administering to a mammalian subject a therapeutically effective amount of
the polypeptide of claim 11 or the polynucleotide of claim 1.

18. A method of diagnosing a pathological condition or a susceptibility to a
25 pathological condition in a subject comprising:

- (a) determining the presence or absence of a mutation in the polynucleotide of
claim 1; and
- (b) diagnosing a pathological condition or a susceptibility to a pathological
30 condition based on the presence or absence of said mutation.

19. A method of diagnosing a pathological condition or a susceptibility to a
pathological condition in a subject comprising:

- (a) determining the presence or amount of expression of the polypeptide of
35 claim 11 in a biological sample; and
- (b) diagnosing a pathological condition or a susceptibility to a pathological
condition based on the presence or amount of expression of the polypeptide.

20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:

- 5 (a) contacting the polypeptide of claim 11 with a binding partner; and
(b) determining whether the binding partner effects an activity of the polypeptide.

21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.

10 22. A method of identifying an activity in a biological assay, wherein the method comprises:

- (a) expressing SEQ ID NO:X in a cell;
(b) isolating the supernatant;
(c) detecting an activity in a biological assay; and
15 (d) identifying the protein in the supernatant having the activity.

23. The product produced by the method of claim 22.

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.